

From: Hutzell, Paula  
Sent: Wednesday, January 02, 2002 4:16 PM  
To: STIC-Biotech/ChemLib; Helmer, Georgia  
Subject: FW: RUSH on FW: sequence search for 09/289346

approved

-----Original Message-----

From: Helmer, Georgia  
Sent: Wednesday, January 02, 2002 4:08 PM  
To: Hutzell, Paula  
Subject: RUSH on FW: sequence search for 09/289346

Paula, I need this search as a Rush because it is due the next biweek. Thanks, Georgia

-----Original Message-----

From: Bui, Phuong  
Sent: Wednesday, January 02, 2002 4:02 PM  
To: Helmer, Georgia  
Subject: RE: sequence search for 09/289346

It's fine. Send it to Paula as a RUSH.

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1 1E01 TEL: 308-3534

12C14

-----Original Message-----

From: Helmer, Georgia  
Sent: Wednesday, January 02, 2002 3:30 PM  
To: Bui, Phuong  
Subject: sequence search for 09/289346

[Bui, Phuong] Would [Bui, Phuong] you please do a search for the following sequences from application # 09/289346:

DNA encoding SEQ ID 2 - 10;  
and polypeptides SEQ ID 2-10. *removed*

Please do an interference search of DNA encoding SEQ ID 2 - 10  
and polypeptides SEQ ID 2-10.

Thanks in advance,  
Georgia L. Helmer Ph.D.  
Patent Examiner  
Crystal Mall 1, 9D14  
AU 1638  
703-308-7023  
Georgia.Helmer@USPTO.gov

FOR OFFICIAL USE ONLY

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 1/3  
Date Completed: 1/4  
Searcher Prep/Review: 10  
Clerical: 12  
Online time: 10

TYPE OF SEARCH:  
NA Sequences: 10  
AA Sequences: 10  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST(where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: as  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

Call 1-800-TEL-308-2234  
Technical Info. Services  
Toby Port  
Dept. of Commerce

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:35:03 ; Search time 144.17 Seconds  
(without alignments)  
35.965 Million cell updates/sec

Title: US-09-289-346A-2  
Perfect score: 362  
Sequence: 1 TLVWGEFQDGAAGGCOT.....FQHNLNSLDRIFDKTPPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	362	100.0	70	AA18678	Mutant peptide der
2	349	96.4	70	AA18677	Peptide fragment f
3	349	96.4	356	AA18687	Amino acid sequenc
4	341	94.2	70	AA18685	Mutant peptide der
5	338	93.4	70	AA18688	Mutant peptide der
6	337	93.1	70	AA18692	Mutant peptide der
7	335	92.5	70	AA18684	Mutant peptide der
8	335	92.5	70	AA18690	Mutant peptide der
9	334	92.3	70	AA18686	Mutant peptide der
10	333	92.0	70	AA18689	Mutant peptide der
11	331	91.4	70	AA18680	Mutant peptide der

12	331	91.4	70	21	AA18691	Mutant peptide der
13	328	90.6	70	21	AA18681	Mutant peptide der
14	327	90.3	70	21	AA18683	Mutant peptide der
15	325	89.8	70	21	AA18682	Mutant peptide der
16	317	87.6	70	21	AA18679	Mutant peptide der
17	215	59.4	353	18	AAW34338	Bean golden mosaic
18	215	59.4	353	18	AAW34332	Bean golden mosaic
19	215	59.4	353	18	AAW34333	Bean golden mosaic
20	215	59.4	353	18	AAW34334	Bean golden mosaic
21	215	59.4	353	18	AAW34335	Bean golden mosaic
22	215	59.4	359	17	AAW88870	Sardinian tomato y
23	215	59.4	359	17	AAW88871	Sardinian tomato y
24	215	59.4	359	17	AAW88872	Sardinian tomato y
25	213	58.8	353	8	AAW70407	ORF 4 gene product
26	213	58.8	361	18	AAW34336	Tomato mottle viru
27	213	58.8	361	18	AAW34324	Tomato mottle viru
28	213	58.8	361	18	AAW34325	Tomato mottle viru
29	213	58.8	361	18	AAW34326	Tomato mottle viru
30	206.5	57.0	361	8	AAW70562	Product of ORF 4 f
31	201	55.5	362	19	AAW56495	Tobacco leaf curli
32	200	55.2	357	18	AAW34329	Tomato yellow leaf
33	200	55.2	357	18	AAW34330	Tomato yellow leaf
34	200	55.2	357	18	AAW34331	Tomato yellow leaf
35	192	53.0	357	18	AAW34337	Tomato yellow leaf
36	69.5	19.2	512	19	AAW68473	HTV-1 strain YBF30
37	68.5	18.9	1693	21	AAW48457	Human laminin 5 po
38	68.5	18.9	1693	21	AAW48459	Human laminin 5 po
39	68.5	18.9	1713	16	AAW70148	Deduced sequence o
40	68.5	18.9	1713	21	AAW48458	Human laminin 5 po
41	68.5	18.9	1724	21	AAW48456	Human laminin 5 po
42	65.5	18.1	619	13	AAW27651	Human calcium chan
43	65.5	18.1	2161	14	AAW33545	Sequence of the al
44	65.5	18.1	2161	16	AAW71001	Human neuronal cal
45	65.5	18.1	2161	21	AAW10568	Human calcium chan

## ALIGNMENTS

RESULT 1.  
AA18678  
ID AAB18678 standard; peptide; 70 AA.  
AC AAB18678;  
XX  
XX  
DT 22-JAN-2001 (first entry)  
XX  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (All) protein.  
KW Geminivirus; replication protein; Rep protein; All; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
FH Key Location/Qualifiers  
FT Misc-difference 12 /note= "wild type residue replaced with Ala"  
FT Misc-difference 13 /note= "wild type residue replaced with Ala"  
FT Misc-difference 15 /note= "wild type residue replaced with Ala"  
XX  
XX  
WO200054573-A1.  
XX  
PN 21-SEP-2000.  
XX  
PD 15-MAR-2000; 2000WO-US06759.  
PF  
XX  
XX  
PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX  
PA (UYN-) UNIV NORTH CAROLINA STATE.





XX  
DR WPI: 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 XX  
 XX  
 PS Disclosure; Page 48; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, bean dwarf mosaic virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;

Query Match 93.4%; Score 338; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 2.5e-34;  
 Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVGDGAAAGCOTSDNDAAEALNASSKEEALQIREKIPEKYLFOFHNLNSL 60  
 Db 1 tlvwgefvdgrsarggcaasdaaaalnaskeaalqirekipekylfqfhlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 6  
 AAB18692  
 ID AAB18692 standard; peptide; 70 AA.  
 AC AAB18692;  
 XX  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 DE  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 XX Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 66  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 69  
 FT /note= "wild type residue replaced with Ala"  
 XX  
 XX WO200054573-A1.  
 XX  
 XX 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR  
 XX 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYN-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.  
 DR Transgenic plants with increased resistance to geminivirus infection  
 XX comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure; Page 50; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, bean dwarf mosaic virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 Query Match 93.1%; Score 337; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 3.4e-34;  
 Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVGDGAAAGCOTSDNDAAEALNASSKEEALQIREKIPEKYLFOFHNLNSL 60  
 Db 1 tlvwgefvdgrsarggcaasdaaaalnaskeaalqirekipekylfqfhlnsl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 7  
 AAB18684  
 ID AAB18684 standard; peptide; 70 AA.  
 AC AAB18684;  
 XX  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 DE  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 XX Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 7  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 8  
 FT /note= "wild type residue replaced with Ala"  
 XX  
 XX WO200054573-A1.  
 XX  
 XX 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR  
 XX 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYN-) UNIV NORTH CAROLINA STATE.  
 PA

```

XX  Hanley-Bowdoin L. Orozco BM, Kong L;
PI  WPI; 2000-618851/59.
XX
DR  Transgenic plants with increased resistance to geminivirus infection
PT  comprise a nucleic acid construct containing a nucleic acid sequence
PT  encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX  Claim 52; Page 45; 73pp; English.
PS
XX  The present sequence represents a mutant peptide, derived from a
CC  geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC  double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC  DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC  protein are used to produce transgenic plants. The mutation in AL1 is
CC  present in a ribosome binding region, and expression of mutant AL1
CC  protein imparts increased resistance to geminivirus infection in the
CC  plant. Mutant AL1 proteins are useful for producing plants having
CC  increased resistance or reduced sensitivity to a geminivirus such as
CC  tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC  virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC  cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC  virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC  virus, cotton leaf curl virus or beet curly top virus.
XX
SQ  Sequence 70 AA;

Query Match 92.5%; Score 335; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 6e-34;
Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 TLVWGEFQVGDGAAGCGQTSDNDAAEALNASSKEEALQIREKIPEKYLFGFHNLSNL 60
    ||||| ||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1 tlwgeaavdgrrsggcgqtsndaaaealnasskeaalqirekipekylfgfhnlnsl 60

QY  61 DRIFDKTPEP 70
    ||||| |||||
Db  61 drifdktp 70

RESULT 8
AAB18690
ID  AAB18690 standard; peptide; 70 AA.
XX
AC  AAB18690;
XX
DT  22-JAN-2001 (first entry)
XX
DE  Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX  Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW  ribosome binding region; resistance; geminivirus infection.
XX
OS  Synthetic.
XX  Tomato golden mosaic virus.
XX
FH  Key Location/Qualifiers
FT  Misc-difference 27 /note= "wild type residue replaced with Ala"
FT  Misc-difference 30 /note= "wild type residue replaced with Ala"
FT  Misc-difference 30 /note= "wild type residue replaced with Ala"
XX
PN  WO200054573-A1.
XX
PD  21-SEP-2000.
XX
XX  15-MAR-2000; 2000WO-US06759.
XX
XX  18-MAR-1999; 99US-0125004.
PR  09-APR-1999; 99US-0289346.

```

---

```

XX  (UUNC-) UNIV NORTH CAROLINA STATE.
PA
XX  Hanley-Bowdoin L. Orozco BM, Kong L;
PI  WPI; 2000-618851/59.
XX
DR  Transgenic plants with increased resistance to geminivirus infection
PT  comprise a nucleic acid construct containing a nucleic acid sequence
PT  encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX  Disclosure; Page 49; 73pp; English.
PS
XX  The present sequence represents a mutant peptide, derived from a
CC  geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC  double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC  DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC  protein are used to produce transgenic plants. The mutation in AL1 is
CC  present in a ribosome binding region, and expression of mutant AL1
CC  protein imparts increased resistance to geminivirus infection in the
CC  plant. Mutant AL1 proteins are useful for producing plants having
CC  increased resistance or reduced sensitivity to a geminivirus such as
CC  tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC  virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC  cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC  virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC  virus, cotton leaf curl virus or beet curly top virus.
XX
SQ  Sequence 70 AA;

Query Match 92.5%; Score 335; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 6e-34;
Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 TLVWGEFQVGDGAAGCGQTSDNDAAEALNASSKEEALQIREKIPEKYLFGFHNLSNL 60
    ||||| ||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1 tlwgeaavdgrrsggcgqtsndaaaealnasskeaalqirekipekylfgfhnlnsl 60

QY  61 DRIFDKTPEP 70
    ||||| |||||
Db  61 drifdktp 70

RESULT 9
AAB18686
ID  AAB18686 standard; peptide; 70 AA.
XX
AC  AAB18686;
XX
DT  22-JAN-2001 (first entry)
XX
DE  Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX  Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW  ribosome binding region; resistance; geminivirus infection.
XX
OS  Synthetic.
XX  Tomato golden mosaic virus.
XX
FH  Key Location/Qualifiers
FT  Misc-difference 24 /note= "wild type residue replaced with Leu"
FT  Misc-difference 25 /note= "wild type residue replaced with Leu"
FT  Misc-difference 26 /note= "wild type residue replaced with Leu"
XX
PN  WO200054573-A1.
XX
PD  21-SEP-2000.
XX

```

```

PF 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant ALI protein with a mutation in the Rb binding region
XX
XX Claim 53; Page 46; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as ALI. ALI binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the ALI
XX protein are used to produce transgenic plants. The mutation in ALI is
XX present in a ribosome binding region, and expression of mutant ALI
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant ALI proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA;
SQ
Query Match 92.3%; Score 334; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 7.9e-34;
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLVWGFEQVDDGAAGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSL 60
Db 1 tlwvgefvdgrsarggcqtsdnllealnasskeaalqirekipekylfgfhlnsl 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70
RESULT 10
AAB18689
ID AAB18689 standard; peptide: 70 AA.
XX
XX AAB18689;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.
XX
XX Geminivirus; replication protein; Rep protein; ALI; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 22 /note= "wild type residue replaced with Ala"
XX
XX Misc-difference 23 /note= "wild type residue replaced with Ala"
XX
XX W0200054573-A1.
XX

```

```

PD 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant ALI protein with a mutation in the Rb binding region
XX
XX Disclosure; Page 48-49; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as ALI. ALI binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the ALI
XX protein are used to produce transgenic plants. The mutation in ALI is
XX present in a ribosome binding region, and expression of mutant ALI
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant ALI proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA;
SQ
Query Match 92.0%; Score 333; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 1.1e-33;
Matches 65; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 TLVWGFEQVDDGAAGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSL 60
Db 1 tlwvgefvdgrsarggcqtsaaaaaalnasskeaalqirekipekylfgfhlnsl 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70
RESULT 11
AAB18680
ID AAB18680 standard; peptide: 70 AA.
XX
XX AAB18680;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.
XX
XX Geminivirus; replication protein; Rep protein; ALI; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 42 /note= "wild type residue replaced with Ala"
XX
XX Misc-difference 43 /note= "wild type residue replaced with Ala"
XX
XX Misc-difference 44

```

```

FT XX /note= "wild type residue replaced with Ala"
PN XX
XX XX
XX XX
PD 21-SEP-2000.
XX XX
XX 15-MAR-2000; 2000WO-US06759.
XX XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI; 2000-618851/59.
XX XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
XX
XX Claim 52; Page 43-44; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA;
SQ
Query Match 91.4%; Score 331; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1.9e-33;
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDAAGGCGTSDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 1 tlvwgefvdgrsarggqtsndaaaealnasskeaalqilaaapekylfghnl 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70
RESULT 12
AAB18691
ID AAB18691 standard; peptide; 70 AA.
XX
XX AAB18691;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 34

```

```

FT Misc-difference 35 /note= "wild type residue replaced with Ala"
FT FT
FT Misc-difference 36 /note= "wild type residue replaced with Ala"
FT FT
FT /note= "wild type residue replaced with Ala"
XX
XX WO200054573-A1.
XX PN
XX 21-SEP-2000.
XX PD
XX
XX 15-MAR-2000; 2000WO-US06759.
XX PF
XX 18-MAR-1999; 99US-0125004.
XX PR
XX 09-APR-1999; 99US-0289346.
XX PR
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX PA
XX
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX PI
XX WPI; 2000-618851/59.
XX DR
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT
PT
XX
XX Disclosure; Page 49; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX Sequence 70 AA;
SQ
Query Match 91.4%; Score 331; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1.9e-33;
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDAAGGCGTSDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 1 tlvwgefvdgrsarggqtsndaaaealnasskeaalqilaaapekylfghnl 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70
RESULT 13
AAB18681
ID AAB18681 standard; peptide; 70 AA.
XX
XX AAB18681;
XX
XX 22-JAN-2001 (first entry)
XX
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
XX
XX

```

OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 47 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 48 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 49 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 PN  
 XX 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 XX  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYN-) UNIV NORTH CAROLINA STATE.  
 XX  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant All protein with a mutation in the Rb binding region  
 PT  
 XX  
 XX Claim 52; Page 44; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as All. All binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the All  
 CC protein are used to produce transgenic plants. The mutation in All is  
 CC present in a ribosome binding region, and expression of mutant All  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant All proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 XX Sequence 70 AA;  
 SQ  
 Query Match 90.6%; Score 328; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 4.4e-33;  
 Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDDGAAAGCGCOTSDNDAAEALNASSKEALQIIRKIPKYLFOFHNLNSNL 60  
 Db 1 tlwgefvdgrsargcgctsdnaaaalnasskealqiirekipaaalfqfhlnslnl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70  
 RESULT 14  
 AAB18683  
 ID AAB18683 standard; peptide; 70 AA.  
 XX  
 XX AAB18683;  
 XX  
 XX 22-JAN-2001 (first entry)  
 DT  
 XX Mutant peptide derived from amino acids 110-179 of Rep (All) protein.  
 DE  
 XX

KW Geminivirus; replication protein; Rep protein; All; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 59 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 61 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 62 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 PN  
 XX 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 XX  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYN-) UNIV NORTH CAROLINA STATE.  
 XX  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI: 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant All protein with a mutation in the Rb binding region  
 PT  
 XX  
 XX Claim 53; Page 45; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as All. All binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the All  
 CC protein are used to produce transgenic plants. The mutation in All is  
 CC present in a ribosome binding region, and expression of mutant All  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant All proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 XX Sequence 70 AA;  
 SQ  
 Query Match 90.3%; Score 327; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 5.8e-33;  
 Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDDGAAAGCGCOTSDNDAAEALNASSKEALQIIRKIPKYLFOFHNLNSNL 60  
 Db 1 tlwgefvdgrsargcgctsdnaaaalnasskealqiirekipkylfghlnslnl 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 aaifdktp 70  
 RESULT 15  
 AAB18682  
 ID AAB18682 standard; peptide; 70 AA.  
 XX  
 XX AAB18682;  
 XX

Job time: 153 sec

DT 22-JAN-2001 (first entry)  
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
DE  
XX  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 52 /note= "wild type residue replaced with Ala"  
FT Misc-difference 54 /note= "wild type residue replaced with Ala"  
FT /note= "wild type residue replaced with Ala"  
FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
FT /note= "wild type residue replaced with Ala"  
XX  
PN WO200054573-A1.  
XX  
XX  
PD 21-SEP-2000.  
XX  
XX  
PF 15-MAR-2000; 2000WO-US06759.  
XX  
XX  
PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX  
PA (UYN-C) UNIV NORTH CAROLINA STATE.  
XX  
XX  
PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX  
DR WPI; 2000-618851/59.  
XX  
XX  
PT Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
XX  
PS Claim 53; Page 44-45; 73pp; English.  
XX  
XX  
CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX  
SQ Sequence 70 AA:

Query Match 89.8%; Score 325; DB 21; Length 70;  
Best Local Similarity 91.4%; Pred. No. 1e-32;  
Matches 64; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVGDGAAGGCGTSDAAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
Db 1 TLVWGEFQVGDGRSARGGCGTSDAAAEALNASSKEEALQIIREKIPKYLFOFHNLSNL 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:35:04 ; Search time 72.79 seconds  
(without alignments)  
73.255 Million cell updates/sec

Title: US-09-289-346A-2  
Perfect score: 362  
Sequence: 1 TLVWGEFQVDAAGGCQT.....FOFHNLSNLDRIFDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	96.4	352	1 QOCVLI	ALI protein - toma
2	249	68.8	361	1 QOCVPT	ALI protein - toma
3	236	65.2	358	2 S07594	hypothetical prote
4	229	63.3	362	1 J01887	ALI protein - toma
5	222	61.3	359	2 S39211	gene C1 protein -
6	218	60.2	349	2 J02300	replicase - pepper
7	218	60.2	349	2 S31875	ALI protein - pepp
8	215	59.4	359	2 S22593	hypothetical prote
9	212	58.6	351	2 J02327	ALI protein - indi
10	212	58.6	355	1 QOCVMI	AVI protein - abut
11	209	57.7	358	1 J01870	ALI protein - toma
12	209	57.7	359	2 S39235	gene C1 protein -
13	209	57.7	385	2 S28360	ALI protein - beet
14	202	55.8	360	2 S39885	replication-associ
15	200	55.2	357	1 QOCVCI	ALI protein - toma
16	142	39.2	131	2 S45059	ACI protein (clone
17	124	34.3	347	1 QOCVSI	ALI protein - squa
18	68.5	18.9	1713	2 A55347	adhesive ligand ep
19	65.5	18.1	1610	2 A46227	voltage-dependent
20	65.5	18.1	1646	2 JH0422	voltage-dependent
21	65.5	18.1	2161	2 JH0564	calcium channel al
22	65.5	18.1	2181	2 A38198	calcium channel al
23	65.5	18.1	2203	2 T42742	voltage-dependent
24	65	18.0	371	2 B69502	alcohol dehydrogen
25	64.5	17.8	299	2 B71967	probable peptidyl-
26	64.5	17.8	335	1 DEBSGF	glyceralddehyde-3-p
27	64.5	17.8	557	2 A47162	thioesterase B (E
28	63.5	17.5	392	2 T45290	acetyl-CoA C-acety
29	63	17.4	397	2 B71078	probable NADH oxid

MG223 homolog F10-  
C1 protein - tobac  
UDP-glucose 4-epim  
CYC1/CYP3 transcri  
probable phosphoe  
acetyl-CoA C-acety  
Fc gamma (1g) rec  
hypothetical prote  
probable flagellar  
E2 protein - human  
CDP-diacylglycerol  
5-aminolevulinate  
ribosomal protein  
hypothetical prote  
hypothetical prote  
cell binding facto

ALIGNMENTS

RESULT 1  
QOCVLI  
ALI protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
C:Accession: A04170  
R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBO J. 3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 96.4%; Score 349; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 1.2e-30;  
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 TLVWGEFQVDAAGGCQTSNDAAAEALNASKEALQIIREKIPEKYLFQFHNLSNL 60  
Db 111 TLVWGEFQVDAAGGCQTSNDAAAEALNASKEALQIIREKIPEKYLFQFHNLSNL 170  
Qy 61 DRIFDKTPPEP 70  
Db 171 DRIFDKTPPEP 180

RESULT 2  
QOCVPT  
ALI protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUID:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-361 <COU>  
A:Cross-references: GB:D00940; NID:g222458; PIDN:BA00782.1; PID:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus ALI protein

```
Query Match      68.8%; Score 249; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 1.1e-19;
Matches 46; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDDGAAAGCGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSNL 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TIEMGLFOIDGRSARGGQQTVDNDAAEALNASTGTEAKMIKEKLPKELFQYHNLSNCL 169

QY 61 DRIFDKTPE 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 170 DRIFDKTPE 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C:Species: cassava latent virus
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: S07594
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07590; MUID:90174930
A:Accession: S07594
A:Status: translation not shown
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <NOR>
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C:Genetics:
C:Map position: segment DNA1
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      65.2%; Score 236; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 2.9e-18;
Matches 42; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDDGAAAGCGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSNL 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 TVEWGFOIDGRSARGGQGSANDAYAKALNSGSKSEALVIRELPKDFVLFQHNLSNL 168

QY 61 DRIFDKTPEP 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 169 DRIFQEPAP 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: C1 protein
C:Species: tomato yellow leaf curl virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: JQ1887
R:Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: JQ1885; MUID:93139778
A:Accession: JQ1887
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DRY>
A:Cross-references: GB:S53251
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      63.3%; Score 229; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 1.7e-17;
Matches 45; Conservative 10; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDDGAAAGCGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSNL 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TIEMGLFOIDGRSARGGQQTVDNDAAEALNASTGTEAKMIKEKLPKELFQYHNLSNCL 169

QY 61 DRIFDKTPEP 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 170 NRIFQTPPEP 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
S39211
gene C1 protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C:Accession: S39211
R:Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A:Description: High similarity among the tomato yellow leaf curl virus isolates from
A:Reference number: S39209
A:Accession: S39211
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <NOR>
A:Cross-references: EMBL:225751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      61.3%; Score 222; DB 2; Length 359;
Best Local Similarity 58.0%; Pred. No. 9.8e-17;
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDDGAAAGCGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSNL 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 LEWGTFOIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYTLHFHNINSNL 170

QY 62 RIFDKTPEP 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 RVFQVPPAP 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
JQ2300
replicase - pepper huasteco virus (component A)
N:Alternate names: ORF AL1 protein
C:Species: pepper huasteco virus
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: JQ2300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: JQ2299; MUID:94015007
A:Accession: JQ2300
A:Molecule type: DNA
A:Residues: 1-349 <TOR>
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      60.2%; Score 218; DB 2; Length 349;
Best Local Similarity 57.1%; Pred. No. 2.6e-16;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDDGAAAGCGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSNL 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TVEWGEFQIDGRSARGGQGSANDTYAKALNSASAEALQIKEEQPOHFFLFQHNIVSNA 169

QY 61 DRIFDKTPEP 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 170 NRIFQTPPEP 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
S31875
AL1 protein - pepper rizado amarillo virus
C:Species: pepper rizado amarillo virus
```

A;Molecule type: DNA

Query Match 57.7%; Score 209; DB 1; Length 358;  
Best Local Similarity .52.9%; Pred. No. 2.6e-15;  
Matches 37; Conservative 17; Mismatches 16; Indels





---

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	349	96.4		352	1	VALL1TGMV	P03567 tomato gold
2	249	68.8		361	1	VALL1PMVV	P27258 potato yell
3	236	65.2		358	1	VALL1CLVK	P14982 cassava lat
4	236	65.2		358	1	VALL1CLVN	P14972 cassava lat
5	229	63.3		362	1	VALL1TYLCA	P36279 tomato yell
6	222	61.3		359	1	VALL1TYLCU	P38609 tomato yell
7	218	60.2		349	1	VALL1PHUV	P06923 pepper huas
8	215	59.4		359	1	VALL1TYLCM	P27260 tomato yell
9	213	58.8		353	1	VALL1BGMV	P05175 bean golden
10	212	58.6		355	1	VALL1BMMW	P21947 abutilon mo
11	209	57.7		358	1	VALL1BCTV	P14991 beet curly
12	209	57.7		361	1	VALL1TMOV	Q06657 tomato mott
13	200	55.2		357	1	VALL1TYLVC	P27259 tomato yell
14	124	34.3		347	1	VALL1SLCV	P29048 squash leaf
15	68.5	18.9		1713	1	LMA3_HUMAN	Q16787 homo sapien
16	65.5	18.1		1610	1	CCAD_HESAU	Q93244 mesocricetu
17	65.5	18.1		2161	1	CCAD_HUMAN	Q01668 homo sapien
18	65.5	18.1		2203	1	CCAD_HAT	P27732 rattus norv
19	64.5	17.8		299	1	Y175_HELPU	Q92mg7 helicobacte
20	64.5	17.8		334	1	G3P_BACST	P00362 bacillus st
21	64.5	17.8		335	1	G3P_BACCO	P15115 bacillus co
22	64.5	17.8		557	1	SASB_ANAPL	Q04791 anas platyr
23	62.5	17.3		419	1	Y223_MYCPN	P75465 mycoplasma
24	62	17.1		295	1	VALL1TYDVA	P31617 tobacco yel
25	62	17.1		1483	1	CYP1_YEAST	P12351 saccharomyc
26	60.5	16.7		136	1	Y452_CAEEL	Q62250 caenorhabdi
27	60	16.6		387	1	VE2_HPV41	P27552 human papil
28	60	16.6		447	1	CDSA_DROME	P56079 d phosphati
29	60	16.6		513	1	HENO_CHICK	P18080 gallus gall
30	59.5	16.4		863	1	AMPN_CAUCR	P37893 caulobacter
31	59	16.3		129	1	RK12_PORPU	P51339 porphyra pu
32	59	16.3		247	1	YCP4_YEAST	P25349 saccharomyc
33	58.5	16.2		299	1	Y175_HELPU	P56112 helicobacte

```

ID VAL1_PYMVV          STANDARD:      PRPT;   361 AA.
P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE ALL PROTEIN.
OS potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
NCBI_TaxID=10828;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91311403; PubMed=1856690;
RX Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.:
RA "The nucleotide sequence of the infectious cloned DNA components of
RT potato yellow mosaic virus."
RL J. Gen. Virol. 72:1515-1520(1991).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; D00940; BAA00782.1; -.
DR PIR; JU0364; QOCVPT.
DR InterPro; IPR001191; Gemin1_AL1.
DR Pfam; PF00799; Gemin1_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemin1_AL1; 1.
DR ATP-binding.
DR NE_BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

```

```

Query Match          58.8%  Score 249;  DB 1;  Length 361;
Best Local Similarity 66.7%;  Pred. No. 4.6e-20;
Matches 46;  Conservative 12;  Mismatches 11;  Indels 0;  Gaps 0;

Qy  1  TLVGEFQVDCAAAGGQTSNDAAALNASSKEEALQIIREKIPEKYLFOFHNLNSNL 60
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  110  TIEWLGFQIDGRSAGQQTVDNAAALNASSGTTKEAMKIIKEKLPEKFLFOYHNLSCNL 169
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy  61  DRIFDKTPE 69
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  170  DRIFMKAPE 178
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 3
VAL1_CLVK
ID  VAL1_CLVK  STANDARD;  PRT;  358 AA.
AC  P14982;
DT  01-APR-1990 (Rel. 14, Created)
DT  01-APR-1990 (Rel. 14, Last sequence update)
DT  01-JUN-1994 (Rel. 29, Last annotation update)
DE  AL1 PROTEIN (40.4 KDA PROTEIN).
```

GN AC1.  
OS Cassava latent virus (strain West Kenya 844).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10818;  
RN [1]  
SEQUENCE FROM N.A.  
RP Stanley J., Gay M.R.;  
RA "Nucleotide sequence of cassava latent virus DNA.";  
RT Nature 301:260-262(1983).  
RL  
CC 1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use.

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR InterPro: IPR001191; Gemini_All.
DR Pfam; PF00799; Gemini_All; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_All; 1.
DR ATP-binding.
KW NP_BIND 220 227 ATP (POTENTIAL).
FT SEQUENCE 358 AA; 40346 MW; EDI73E75EE92D69 CRC64;
SQ
Query Match 65.2%; Score 236; DB 1; Length 358;

```

	Query Match	65.2%	Score 236;	DB 1:	Length 358;
	Best Local Similarity	60.0%;	Pred. No. 1.2e-18;		
	Matches	42;	Conservative 14;	Mismatches 14;	Indels 0; Gaps 0;
Qy	1	TLVNGETQVDGAAAGCGQTNDAAAALNASSKEEALQITIREKIPKYLFQFHNLNSNL	60		
		: :			
Dd	109	TVEWQGQIDGRSGRGQQSANDAYAKALNSGSKSEALNVIRELVPPDKVFQFHNLSNL	168		
Qy	61	DRIFDKTPEP	70	- - - - -	
		: : : : : : :			
Dd	169	DRIFQEPPAP	178	- - - - -	

RESULT	4	
VAL1_CLVN		
ID	VAL1_CLVN	STANDARD; PRT; 358 AA.
AC	P14972;	
DT	01-APR-1990 (Rel. 14, Created)	
DT	01-APR-1990 (Rel. 14, Last sequence update)	
DT	01-JUN-1994 (Rel. 29, Last annotation update)	
DE	ALI PROTEIN (40.4 KDA PROTEIN).	
DE	ALI	
GN	AC1.	
OS	Cassava latent virus (strain Nigerian).	
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.	
NCBI_TaxID=10819;		
RN	[1]	

SEQUENCE FROM N.A.  
MEDLINE=90174930; PubMed=2308831;  
MORRIS B., Coates L., Lowe S., Richardson K., Eddy P.;  
RT "Nucleotide sequence of the infectious cloned DNA components of  
RT African cassava mosaic virus (Nigerian strain).";  
RL Nucleic Acids Res. 18:197-198(1990).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC		EMBL; X17095; CAA34953.l; -.	
DR	DR	PIR; S07594; S07594.	
DR	DR	InPro; IPR001191; Gemini_ALL.	
DR	DR	Pfam; PF00799; Gemini_AL1; 1.	
DR	DR	PRINTS; PR00227; GEMCOATALL.	
DR	DR	PRINTS; PR00228; GEMCOATCLVL1.	
DR	DR	Prodrom; PD000736; Gemini_AL1; 1.	
KW	ATP-binding.	220 227 ATP (POTENTIAL).	
FT	NE_BIND	358 AA; 40435 MW; 1DB16BBOCB2D5E2C CRC64;	
SQ	SEQUENCE		

Query Match	65.2%	Score 236;	DB 1;	Length 358;
Best Local Similarity	60.0%;	Pred. No. 1.2e-18;		
Matches 42;	Conservative	14;	Mismatches 14;	Indels
				Caps 0;





[illegible]

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR PIR; A36214; QOCVM1
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLL.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding. 221 228 ATP (POTENTIAL).
FT NP_BIND
SQ SEQUENCE 355 AA; 40257 MW; 16A2CA8A63251E95 CRC64;

Query Match 58.6%; Score 212; DB 1; Length 355;
Best Local Similarity 55.7%; Pred. No. 5.1e-16;
Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGGAAAGCGCOTSDAAAEALNASSKEALQIIREKIPEKYLFOFHNLNSNL 60
Db 110 TAEWGEFQIDGRSARGGOQTANDSYAKALNAGDVQSALNILEQPKDYVLQHNHNSNL 169

QY 61 DRIFDKTPEP 70
Db 170 ERIFAKAPEP 179

RESULT 11
VAL1_BCTV
ID VAL1_BCTV STANDARD; PRT; 358 AA.
AC P14991;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE ALL PROTEIN (40.8 KDA PROTEIN).
OS Beet curly top virus (BCTV).
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX NCBI_TaxID=10840;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT "The nucleotide sequence of an infectious clone of the geminivirus
RT beet curly top virus.";
RL EMBO J. 5:1761-1767(1986).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLL.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding. 222 229 ATP (POTENTIAL).
FT NP_BIND
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 57.7%; Score 209; DB 1; Length 358;
```

```
Best Local Similarity 52.9%; Pred. No. 1.1e-15;
Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGGAAAGCGCOTSDAAAEALNASSKEALQIIREKIPEKYLFOFHNLNSNL 60
Db 110 TAEWGEFQIDGRSARGGOQTANDSYAKALNAGDVQSALNILEQPKDYVLQHNHNSNL 169

QY 61 DRIFDKTPEP 70
Db 170 ERIFAKAPEP 179

RESULT 12
VAL1_TMOV
ID VAL1_TMOV STANDARD; PRT; 361 AA.
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
OS All.
GN Tomato mottle virus (isolate Florida) (TMov).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=36449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107858; PubMed=1469361;
RA Abouzid A.M., Polston J.E., Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT isolated from tomatoes in Florida.";
RL J. Gen. Virol. 73:3225-3229(1992).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14460; AAC32414.1; -
DR PIR; JQ1870; JQ1870.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLL.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding. 222 229 ATP (BY SIMILARITY).
FT NP_BIND
SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEEAC6950 CRC64;

Query Match 57.7%; Score 209; DB 1; Length 361;
Best Local Similarity 52.9%; Pred. No. 1.1e-15;
Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGGAAAGCGCOTSDAAAEALNASSKEALQIIREKIPEKYLFOFHNLNSNL 60
Db 110 TAEWGEFQIDGRSARGGOQTANDSYAKALNAGDVQSALNILEQPKDYVLQHNHNSNL 169

QY 61 DRIFDKTPEP 70
Db 170 ERIFAKAPEP 179

RESULT 13
VAL1_TYLCV
ID VAL1_TYLCV STANDARD; PRT; 357 AA.
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
```

DE ALL PROTEIN (C1 PROTEIN).  
 GN C1.  
 OS Tomato yellow leaf curl virus (TYLCV).  
 OC Viruses; SS DNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10832;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92024070; PubMed=1926771;  
 RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
 RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
 with a single genomic component.";  
 RL Virology 185:151-161(1991).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X15636; CAJ33688.1; -  
 DR PIR; D40779; QOCVCI.  
 DR InterPro; IPR001191; Gemini\_AL1.  
 DR Pfam; PF00799; Gemini\_AL1.  
 DR PRINTS; PR00227; GEMCOATALL.  
 DR PRINTS; PR00228; GEMCOATCLVLL.  
 DR ProDom; PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND 219 226 ATP (POTENTIAL).  
 SQ SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;  
 Query Match 55.2%; Score 200; DB 1; Length 357;  
 Best Local Similarity 62.3%; Pred. No. 1.1e-14;  
 Matches 38; Conservative 11; Mismatches 12; Indels 0; Gaps 0;  
 QY 4 WGEFQVDDGAAAGCCQTSNDAAAEALNASSKEEALQIREKIPKYLFOFHNLSNLDRI 63  
 Db 111 FGVSQIDGRSARGQOOSANDAYAEALNSGSKSEALNILEKAPKDYLOFHNLSNLDRI 170  
 QY 64 F 64  
 Db 171 F 171  
 RESULT 14  
 VAL1\_SLCV STANDARD; PRT; 347 AA.  
 AC P29048;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE ALL PROTEIN.  
 OS Squash leaf curl virus.  
 OC Viruses; SS DNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91082449; PubMed=1984668;  
 RA Lazarowitz S.G., Lazdins I.B.;  
 RT "Infectivity and complete nucleotide sequence of the cloned genomic  
 components of a bipartite squash leaf curl geminivirus with a broad  
 host range phenotype.";  
 RL Virology 180:58-69(1991).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M38183; AAC32410.1; ALT\_INIT.  
 DR PIR; C36785; QOCVSI.  
 DR InterPro; IPR001191; Gemini\_AL1.  
 DR Pfam; PF00799; Gemini\_AL1.  
 DR PRINTS; PR00227; GEMCOATALL.  
 DR PRINTS; PR00228; GEMCOATCLVLL.  
 DR ProDom; PD000736; Gemini\_AL1; 1.  
 KW ATP-binding.  
 FT NP\_BIND 218 225 ATP (POTENTIAL).  
 SQ SEQUENCE 347 AA; 39110 MW; AFDAEBDE122110E CRC64;  
 Query Match 34.3%; Score 124; DB 1; Length 347;  
 Best Local Similarity 37.9%; Pred. No. 2.2e-06;  
 Matches 25; Conservative 14; Mismatches 23; Indels 4; Gaps 1;  
 QY 5 GEFOVDGAAAGCCQTSNDAAAEALNASSKEEALQIREKIPKYLFOFHNLSNLDRI 64  
 Db 116 GQYKVSQ-----GSKSNKDDVYHNAGSAGEALDLIKAGDPKFTFVYHNLLANVERLF 171  
 QY 65 DKTPEP 70  
 Db 172 QKPEP 177  
 RESULT 15  
 LMA3\_HUMAN STANDARD; PRT; 1713 AA.  
 ID Q16787; Q13679; Q13680;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE LAMININ ALPHA-3 CHAIN PRECURSOR (EPILIGRIN 170 KDA SUBUNIT) (E170).  
 GN LMA3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Keratinocytes;  
 RX MEDLINE=94357926; PubMed=8077230;  
 RA Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;  
 RT "Cloning of the Lma3 gene encoding the alpha 3 chain of the adhesive  
 RT ligand epiligrin. Expression in wound repair.";  
 RL J. Biol. Chem. 269:22779-22787(1994).  
 RN [2]  
 RP SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).  
 RX MEDLINE=96163880; PubMed=8586427;  
 RA Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,  
 RA Vitto J., Ortonne J.-P., Meneguzzi G.;  
 RT "Cloning of the laminin alpha 3 chain gene (LMA3) and identification  
 RT of a homozygous deletion in a patient with Herlitz junctional  
 RT epidermolysis bullosa.";  
 RL Genomics 30:273-280(1995).  
 CC -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT, BY INTERACTING  
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
 CC -!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION  
 CC VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-  
 CC 6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE  
 CC PHOSPHORYLATION OF PP125-FAK AND p80, (3) DIFFERENTIATION OF  
 CC KERATINOCYTES.  
 CC -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLYCOLES AT EACH END.  
 CC THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/  
 CC NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)

[illegible]

Query Match 18.9%; Score 68.5; DB 1; Length 1713;  
Best Local Similarity 36.1%; Pred. NO. 15;  
Matches 22; Conservative 11; Mismatches 21; Indels 7





```
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; SSDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWY- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR InterPro; IP001191; Gemin_A1.
DR Pfam; PF00799; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_A1; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 77.6%; Score 281; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 9.2e-23;
Matches 54; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGGQOTVNDAAAEALNASSKEPALQIIREKPKYLFQFHNLSNL 60
Db 111 TVWGEFQVDSRGSSRGQOTVNDAAAEALNAPDKRTALQIIREKPKYLFQFHNLSNL 170
QY 61 DRIFDKTPEP 70
Db 171 DRIFAKAPEP 180

RESULT 3
O9WHF6 PRELIMINARY; PRT; 226 AA.
ID O9WHF6
AC O9WHF6
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; SSDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN96-H5KW;
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M., Maxwell D.P.;
RL "Molecular characterization and DNA-based detection methods for vegetable-infecting geminiviruses in Central America.";
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131071; RAD33471.1; -.
DR InterPro; IP001191; Gemin_A1.
DR Pfam; PF00799; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_A1; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 77.6%; Score 281; DB 12; Length 226;
Best Local Similarity 72.9%; Pred. No. 9.2e-23;
Matches 51; Conservative 13; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGGQOTVNDAAAEALNASSKEPALQIIREKPKYLFQFHNLSNL 60
Db 111 TVWGEFQVDSRGSSRGQOTVNDAAAEALNAPDKRTALQIIREKPKYLFQFHNLSNL 170
```

```
Db 111 TIWGEFQIDGRSARGGQOTDAAAEALNASSKEAMRIKPKELFQYHNLSNL 170
QY 61 DRIFDKTPEP 70
Db 171 DRIFAKAPEP 180

RESULT 4
O67574 PRELIMINARY; PRT; 361 AA.
ID O67574
AC O67574
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE REPLICATIVE PROTEIN.
GN ALL.
OS Bean golden mosaic virus.
OC Viruses; SSDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10839;
RN [1]
RP SEQUENCE FROM N.A.
RC Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G., Maxwell D.P., Russell D.R.;
RA Maxwell D.P., Russell D.R.;
RL Phytopathology 81:980-985(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C., Morales F.J., Maxwell D.P.;
RA Morales F.J., Maxwell D.P.;
RL Plant Dis. 75:336-342(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M88686; AAA46312.1; -.
DR InterPro; IP001191; Gemin_A1.
DR Pfam; PF00799; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_A1; 1.
SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.1%; Score 279; DB 12; Length 361;
Best Local Similarity 76.1%; Pred. No. 2.6e-22;
Matches 51; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 4 WGEFQVDSARGGQOTSDAAAEALNASSKEALQIIREKPKYLFQFHNLSNLDRI 63
Db 113 WGEFQVDSARGGQOTDAAAEALNASSKEAMQIIREKPKELFQYHNLSNLDRI 172
QY 64 FDKTPEP 70
Db 173 FTKAPDP 179

RESULT 5
O9QDB1 PRELIMINARY; PRT; 225 AA.
ID O9QDB1
AC O9QDB1
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS cowpea golden mosaic geminivirus.
OC Viruses; SSDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=69263;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGMV-BR;
RA Faria J.C.;
RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from Brazil.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
```



RESULT  
P88975  
9

ID P88975 PRELIMINARY; PRT; 149 AA.  
 AC P88975; Thesis (1996). Biochemistry, University of the West Indies, Jamaica.  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN AC1.  
 OS Macroptilium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E.;  
 RL Thesis (1996). Biochemistry, University of the West Indies, Jamaica.  
 DR EMBL; 075278; AA836919.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR ProDom; PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 1 149  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 16785 MW; E4CF5EED4C9CD508 CRC64;

Query Match 73.5%; Score 266; DB 12; Length 149;  
 Best Local Similarity 68.6%; Pred. No. 2.4e-21;  
 Matches 48; Conservative 13; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDDGAAAGCGCOTSDAAAEALNASSKEALQIIRKIKPEKYLFOFHNLNSNL 60  
 Db 52 TIWGVFQIDGRSARGCGQOTSDAAAEALNASSKEALQIIRKIKPEKYLFOFHNLNSNL 111  
 QY 61 DRIFDKTPEP 70  
 Db 112 DRIFMKDPEP 121

RESULT 10  
 QY9YLA4 PRELIMINARY; PRT; 233 AA.  
 ID QY9YLA4;  
 AC QY9YLA4;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Macroptilium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.;  
 RL "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses from Jamaica.";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.; McLaughlin W.A.; Maxwell D.P.;  
 RL "Molecular characterization of two distinct geminiviruses infecting M. lathyroides from Jamaica.";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF098940; AAD17850.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR ProDom; PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 233 233  
 SQ SEQUENCE 233 AA; 26355 MW; AA490AF4D2166A02 CRC64;

Query Match 73.5%; Score 266; DB 12; Length 233;  
 Best Local Similarity 68.6%; Pred. No. 4e-21;  
 Matches 48; Conservative 13; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDDGAAAGCGCOTSDAAAEALNASSKEALQIIRKIKPEKYLFOFHNLNSNL 60  
 Db 110 TIWGVFQIDGRSARGCGQOTSDAAAEALNASSKEALQIIRKIKPEKYLFOFHNLNSNL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 DRIFMKDPEP 179  
 RESULT 11  
 ID Q39180 PRELIMINARY; PRT; 234 AA.  
 AC Q39180;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 OS potato yellow mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10827;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TOMATO STRAIN;  
 RA Guzman P.; Arredondo C.R.; Emmatty D.; Portillo R.J.; Gilbertson R.L.;  
 RL Plant Dis. 81:312-312(1997).  
 DR EMBL; AF026553; AAB82605.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS; PR00227; GEMCOATL1.  
 DR ProDom; PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 234 234  
 SQ SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;

Query Match 71.8%; Score 260; DB 12; Length 234;  
 Best Local Similarity 67.1%; Pred. No. 1.8e-20;  
 Matches 47; Conservative 14; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDDGAAAGCGCOTSDAAAEALNASSKEALQIIRKIKPEKYLFOFHNLNSNL 60  
 Db 110 TIWGVFQIDGRSARGCGQOTSDAAAEALNASSKEALQIIRKIKPEKYLFOFHNLSSDL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 DRIFMKAPNP 179  
 RESULT 12  
 ID Q9Z089 PRELIMINARY; PRT; 190 AA.  
 AC Q9Z089;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA3;  
 RA Ooi K.; Ohshita S.; Ishii I.; Yahara T.;  
 RL "Molecular phylogeny of geminivirus infecting wild plants in Japan.";  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL; AB001315; BAA34033.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.

OS tobacco leaf curl virus,  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Search completed: January 3, 2002, 15:56:24

Fri Jan 4 09:37:40 2002

us-09-289-346a-2.rspt

Page 6

Job time: 1120 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:35:03 ; Search time 65.28 Seconds  
(without alignments)  
24.130 Million cell updates/sec

Title: US-09-289-346A-2

Perfect score: 362

Sequence: 1 TLVWGEFQVDAAGGCQT.....FOPHNLNSLDRFDKTPPE 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	59.4	353	4	US-08-838-151A-44
2	215	59.4	353	4	US-08-838-151A-46
3	215	59.4	353	4	US-08-838-151A-49
4	215	59.4	353	4	US-08-838-151A-52
5	215	59.4	353	4	US-08-838-151A-55
6	215	59.4	359	4	US-08-809-103B-2
7	215	59.4	359	4	US-08-809-103B-4
8	215	59.4	359	4	US-08-809-103B-6
9	215	59.4	359	4	US-08-809-103B-8
10	213	58.8	361	4	US-08-838-151A-2
11	213	58.8	361	4	US-08-838-151A-4
12	213	58.8	361	4	US-08-838-151A-6
13	213	58.8	361	4	US-08-838-151A-8
14	200	55.2	357	4	US-08-838-151A-24
15	200	55.2	357	4	US-08-838-151A-27
16	200	55.2	357	4	US-08-838-151A-30
17	192	53.0	357	4	US-08-838-151A-20
18	68.5	18.9	1713	3	PCT-US94-10261A-24
19	68.5	18.9	1713	3	US-07-745-206A-2
20	65.5	18.1	2161	1	US-08-455-543A-49
21	65.5	18.1	2161	1	US-08-455-543A-51
22	65.5	18.1	2161	2	US-08-223-305C-49
23	65.5	18.1	2161	2	US-08-223-305C-51
24	65.5	18.1	2161	2	US-08-311-363-2
25	65.5	18.1	2161	2	US-08-311-363-2
26	63.5	17.5	334	6	5290690-11
27	60	16.6	446	2	US-08-672-814D-11

28	60	16.6	446	4	US-09-333-696-11	Sequence 11, Appl
29	59	16.3	454	3	US-08-446-100-31	Sequence 31, Appl
30	58.5	16.2	439	1	US-07-637-870-9	Sequence 9, Appl
31	58.5	16.2	439	1	US-07-637-399-6	Sequence 6, Appl
32	58.5	16.2	439	1	US-08-112-703-6	Sequence 6, Appl
33	58	16.0	844	3	US-09-029-267-20	Sequence 20, Appl
34	57	15.7	454	3	US-08-446-100-30	Sequence 30, Appl
35	56	15.5	454	3	US-08-446-100-26	Sequence 26, Appl
36	56	15.5	454	3	US-08-446-100-27	Sequence 27, Appl
37	56	15.5	454	3	US-08-446-100-28	Sequence 28, Appl
38	56	15.5	454	3	US-08-446-100-29	Sequence 29, Appl
39	56	15.5	548	2	US-08-871-266B-17	Sequence 17, Appl
40	56	15.5	548	2	US-09-018-864A-17	Sequence 17, Appl
41	56	15.5	548	3	US-08-871-267B-23	Sequence 23, Appl
42	56	15.5	548	4	US-09-618-419-23	Sequence 23, Appl
43	55.5	15.3	111	1	US-08-185-414E-2	Sequence 2, Appl
44	55.5	15.3	770	1	US-08-445-135-2	Sequence 2, Appl
45	55	15.2	384	4	US-08-928-442-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-08-838-151A-44  
; Sequence 44, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
; TITLE OF INVENTION: Genes  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,151A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-838-151A-44

Query Match 59.4%; Score 215; DB 4; Length 353;  
Best Local Similarity 57.1%; Pred. No. 1.4e-19;  
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

OY 1 TLVWGEFQVDAAGGCQTSNDAAAEALNASSKEALQITREKIPKYLFOFHNLNSNL 60



Query Match 59.4%; Score 215; DB 4; Length 359;  
Best Local Similarity 55.1%; Pred. No. 1.4e-19;  
Matches 38; Conservative 12; Mismatches 19; Indels

Query Match 59.4%; Score 215; DB 4; Length 359;  
Best Local Similarity 55.1%; Pred. No. 1.4e-19;  
Matches 38; Conservative 12; Mismatches 19; Indels









Query Match	55.2%	Score 200;	DB 4;	Length 357;
Best Local Similarity	62.3%;	Pred. No. 1.2e-17;		
Matches 38; Conservative 11;	Mismatches 12;	Indels 0;	Gaps 0;	
OY	4	WGFEVDGAAAGCCQTSNDAAAEALNASKEEALQITIREKIPKYLFQFHNLSNLDRI	63	
	:	:     :                                 :		
Dd	111	FQVSQIDGRARGQQSANDAYAEALNSGGSKSEALNLKERAPKDYLILQFHNLSSNLDRI	170	
OY	64	F 64		
Dd	171	F 171		

Search completed: January 3, 2002, 15:38:48  
Job time: 225 sec

```

RESULT 15
US-08-838-151A-27
: Sequence 27, Application US/08838151A
: Patent No. 6291743
: GENERAL INFORMATION:
: APPLICANT: Stout, John T
: APPLICANT: Luu, Hang T
: APPLICANT: Maxwell, Douglas
: APPLICANT: Ahlquist, Paul
: APPLICANT: Hanson, Steve
: TITLE OF INVENTION: Transgenic pl
: TITLE OF INVENTION: Genes
: NUMBER OF SEQUENCES: 63
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dressler, Rocky, Mil
: STREET: Two Prudential Plaza, S
: CITY: Chicago
: STATE: Illinois
: COUNTRY: U.S.A.
: ZIP: 60601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/838,1
: FILING DATE:
: CLASSIFICATION: 800

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:37:36 ; Search time 144.17 seconds  
(without alignments)  
35.965 Million cell updates/sec

Title: US-09-289-346A-3  
Perfect score: 358  
Sequence: 1 TLWGEFQVDRSGRGCCOT.....FQPHNLNSLDRIFDKTPPE 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	358	100.0	70	AA18679	Mutant peptide der
2	347	96.9	70	AA18691	Mutant peptide der
3	343	95.8	70	AA18690	Mutant peptide der
4	332	92.7	70	AA18677	Peptide fragment f
5	332	92.7	356	AA18687	Amino acid sequenc
6	324	90.5	70	AA18685	Mutant peptide der
7	321	89.7	70	AA18688	Mutant peptide der
8	320	89.4	70	AA18692	Mutant peptide der
9	318	88.8	70	AA18684	Mutant peptide der
10	317	88.5	70	AA18678	Mutant peptide der
11	317	88.5	70	AA18686	Mutant peptide der

12	316	88.3	70	21	AA18689	Mutant peptide der
13	314	87.7	70	21	AA18680	Mutant peptide der
14	311	86.9	70	21	AA18681	Mutant peptide der
15	310	86.6	70	21	AA18683	Mutant peptide der
16	308	86.0	70	21	AA18682	Mutant peptide der
17	218	60.9	361	18	AAW34336	Tomato mottle viru
18	218	60.9	361	18	AAW34324	Tomato mottle viru
19	218	60.9	361	18	AAW34325	Tomato mottle viru
20	218	60.9	361	18	AAW34326	Tomato mottle viru
21	217	60.6	353	18	AAW34338	Bean golden mosaic
22	217	60.6	353	18	AAW34332	Bean golden mosaic
23	217	60.6	353	18	AAW34333	Bean golden mosaic
24	217	60.6	353	18	AAW34334	Bean golden mosaic
25	217	60.6	353	18	AAW34335	Bean golden mosaic
26	215	60.1	353	8	AAW70407	ORF 4 gene product
27	212	59.2	359	17	AAW88870	Sardinian tomato y
28	212	59.2	359	17	AAW88871	Sardinian tomato y
29	212	59.2	359	17	AAW88872	Sardinian tomato y
30	200.5	56.0	361	8	AAW70562	Product of ORF 4 f
31	198	55.3	362	19	AAW56495	Tobacco leaf curl
32	190	53.1	357	18	AAW34337	Tomato yellow leaf
33	190	53.1	357	18	AAW34329	Tomato yellow leaf
34	190	53.1	357	18	AAW34330	Tomato yellow leaf
35	190	53.1	357	18	AAW34331	Tomato yellow leaf
36	73.5	20.5	304	22	AAW39860	Human polypeptide
37	73.5	20.5	406	22	AAW92868	Human protein sequ
38	73.5	20.5	425	22	AAW74314	Human splicing fac
39	73.5	20.5	427	22	AAW41646	Human polypeptide
40	62.5	17.5	579	19	AAW63707	Human hsk2 protein
41	61.5	17.2	66	11	AAW80867	Synthetic antife
42	61	17.0	197	21	AAW24212	Arabidopsis thalia
43	61	17.0	218	21	AAW24211	Arabidopsis thalia
44	61	17.0	1212	20	AAW88287	Rat (Na,K)-ATPase
45	61	17.0	1212	20	AAW88288	Rat (Na,K)-ATPase

## ALIGNMENTS

RESULT 1  
AA18679  
ID AA18679 standard; peptide; 70 AA.  
AC AA18679;  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (ALL) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; ALL; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
FH Key Location/Qualifiers  
FT Misc-difference 27 /note= "wild type residue replaced with Ala"  
FT Misc-difference 30 /note= "wild type residue replaced with Ala"  
FT Misc-difference 34 /note= "wild type residue replaced with Ala"  
FT Misc-difference 35 /note= "wild type residue replaced with Ala"  
FT Misc-difference 36 /note= "wild type residue replaced with Ala"  
XX WO200054573-A1.  
XX 21-SEP-2000.  
XX 15-MAR-2000; 2000WO-US06759.

```
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant ALI protein with a mutation in the Rb binding region
PT
XX
XX Claim 52; Page 43; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as ALI. ALI binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the ALI
CC protein are used to produce transgenic plants. The mutation in ALI is
CC present in a ribosome binding region, and expression of mutant ALI
CC plant. Mutant ALI proteins are useful for producing plants having
CC increased resistance to geminivirus infection in the
CC plant. Mutant ALI proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX Sequence 70 AA;

Query Match 100.0%; Score 358; DB 21; Length 70;
Best Local Similarity 100.0%; Pred. No. 2.8e-39;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIIREKIPEKYLFOFHNLNSNL 60
Db 1 TLVWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIIREKIPEKYLFOFHNLNSNL 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70

RESULT 2
AAB18691
ID AAB18691 standard; peptide; 70 AA.
XX
XX AAB18691;
AC
XX
XX 22-JAN-2001 (first entry)
DE Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.
XX Geminivirus; replication protein; Rep protein; ALI; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
OS Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 34 /note= "wild type residue replaced with Ala"
FT Misc-difference 35 /note= "wild type residue replaced with Ala"
FT Misc-difference 36 /note= "wild type residue replaced with Ala"
FT Misc-difference 36 /note= "wild type residue replaced with Ala"
XX
XX WO200054573-A1.
XX
```

```
PD 21-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US06759.
XX
XX 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
XX (UYNC-) UNIV NORTH CAROLINA STATE.
PA
XX Hanley-Bowdoin L, Orozco BM, Kong L;
PI WPI; 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant ALI protein with a mutation in the Rb binding region
PT
XX
XX Disclosure; Page 49; 73pp; English.
XX
XX The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as ALI. ALI binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the ALI
CC protein are used to produce transgenic plants. The mutation in ALI is
CC present in a ribosome binding region, and expression of mutant ALI
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant ALI proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX Sequence 70 AA;

Query Match 96.9%; Score 347; DB 21; Length 70;
Best Local Similarity 97.1%; Pred. No. 7.7e-38;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIIREKIPEKYLFOFHNLNSNL 60
Db 1 TLVWGEFQVDRSARGCGCOTSDNDAALAAASSAAALQIIREKIPEKYLFOFHNLNSNL 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70

RESULT 3
AAB18690
ID AAB18690 standard; peptide; 70 AA.
XX
XX AAB18690;
AC
XX
XX 22-JAN-2001 (first entry)
DT Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.
DE Geminivirus; replication protein; Rep protein; ALI; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
XX Synthetic.
OS Tomato golden mosaic virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 27 /note= "wild type residue replaced with Ala"
FT Misc-difference 30 /note= "wild type residue replaced with Ala"
FT Misc-difference 30 /note= "wild type residue replaced with Ala"
XX
XX
```

```
PN WO200054573-A1.
XX
XX PD
XX 21-SEP-2000.
XX
XX PF 15-MAR-2000; 2000WO-US06759.
XX
XX PR 18-MAR-1999; 99US-0125004.
XX
XX PR 09-APR-1999; 99US-0289346.
XX
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX DR WPI; 2000-618851/59.
XX
XX PT Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX PS Disclosure; Page 49; 73pp; English.
XX
XX CC The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX protein are used to produce transgenic plants. The mutation in AL1 is
XX present in a ribosome binding region, and expression of mutant AL1
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant AL1 proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX SQ Sequence 70 AA;

Query Match 95.8%; Score 343; DB 21; Length 70;
Best Local Similarity 95.7%; Pred. No. 2.6e-37;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGRSARGCOTSDAAAAAALQIIREKIPEKYLFOFHNLNSNL 60
DB 1 TLVWGEFVDGRSARGCOTSDAAAAAALQIIREKIPEKYLFOFHNLNSNL 60

QY 61 DRIFDKTPEP 70
DB 61 drifdktp 70

RESULT 4
AAB18677
ID AAB18677 standard; peptide; 70 AA.
XX
XX AC AAB18677;
XX
XX 22-JAN-2001 (first entry)
XX
XX DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
XX KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX OS Tomato golden mosaic virus.
XX
XX PN WO200054573-A1.
XX
XX PD 21-SEP-2000.
XX
XX PF 15-MAR-2000; 2000WO-US06759.
XX
```

```
XX 18-MAR-1999; 99US-0125004.
XX
XX PR 09-APR-1999; 99US-0289346.
XX
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
XX DR WPI; 2000-618851/59.
XX
XX PT Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
XX PS Disclosure; Page 18; 73pp; English.
XX
XX CC The present sequence is derived from a geminivirus replication (Rep)
XX protein, which is also known as AL1. AL1 binds double-stranded DNA,
XX catalyses cleavage and ligation of single-stranded DNA, and interacts
XX with other viral and host proteins. Mutants of the AL1 protein are used
XX to produce transgenic plants. The mutation in AL1 is present in a
XX ribosome binding region, and expression of mutant AL1 protein imparts
XX increased resistance to geminivirus infection in the plant. Mutant AL1
XX proteins are useful for producing plants having increased resistance or
XX reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
XX tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
XX virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
XX yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
XX squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
XX beet curly top virus.
XX
XX SQ Sequence 70 AA;

Query Match 92.7%; Score 332; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 7e-36;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGRSARGCOTSDAAAAAALQIIREKIPEKYLFOFHNLNSNL 60
DB 1 TLVWGEFVDGRSARGCOTSDAAAAAALQIIREKIPEKYLFOFHNLNSNL 60

QY 61 DRIFDKTPEP 70
DB 61 drifdktp 70

RESULT 5
AAB18687
ID AAB18687 standard; peptide; 356 AA.
XX
XX AC AAB18687;
XX
XX 22-JAN-2001 (first entry)
XX
XX DE Amino acid sequence of a geminivirus replication protein of TGMV.
XX
XX KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX ribosome binding region; resistance; geminivirus infection.
XX
XX OS Tomato golden mosaic virus.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 354 /note= "unspecified amino acid"
XX
XX PN WO200054573-A1.
XX
XX PD 21-SEP-2000.
XX
XX PF 15-MAR-2000; 2000WO-US06759.
XX
```

PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYN-) UNIV NORTH CAROLINA STATE.  
 PA  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 XX WPI; 2000-618851/59.  
 DR  
 DR  
 XX

Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT

PS Disclosure; Page 47-48; 73pp; English.

CC The present sequence represents a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.

XX Sequence 356 AA;

Query Match 92.7%; Score 332; DB 21; Length 356;  
 Best Local Similarity 92.9%; Pred. No. 5.4e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGCQTSNDAAAALAAASAAAALQIIEKIPKYLFOFHNLNSNL 60  
 Db ||||||| ||||||| ||||||| || ||| ||||||| ||||||| |||||||

QY 110 tlwgefvdgrsarggcqtsndaaacalnasskeaalqlirekipekylfqhlnsl 169  
 Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 6  
 AAB18685  
 ID AAB18685 standard; peptide; 70 AA.  
 XX  
 AC AAB18685;  
 XX

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.  
 OS Tomato golden mosaic virus.

XX Key Location/Qualifiers  
 FH Misc-difference 10

FT /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.

XX 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX

PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYN-) UNIV NORTH CAROLINA STATE.  
 PA  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX  
 XX WPI; 2000-618851/59.  
 DR  
 DR  
 XX

Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT

XX Claim 53; Page 46; 73pp; English.

XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.

XX Sequence 70 AA;

Query Match 90.5%; Score 324; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 7.7e-35;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGCQTSNDAAAALAAASAAAALQIIEKIPKYLFOFHNLNSNL 60  
 Db ||||||| ||||||| ||||||| || ||| ||||||| ||||||| |||||||

QY 61 DRIFDKTPEP 70  
 Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 7  
 AAB18688  
 ID AAB18688 standard; peptide; 70 AA.  
 XX  
 AC AAB18688;  
 XX

DT 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.  
 OS Tomato golden mosaic virus.

XX Key Location/Qualifiers  
 FH Misc-difference 19

FT /note= "wild type residue replaced with Ala"  
 XX  
 PN WO200054573-A1.

XX 21-SEP-2000.

XX





PN WO200054573-A1.  
 PD 21-SEP-2000.  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT .  
 XX Claim 52; Page 45; 73pp; English.  
 PS The present sequence represents a mutant peptide, derived from a  
 XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX Sequence 70 AA:  
 SQ  
 Query Match 88.8%; Score 318; DB 21; Length 70;  
 Best Local Similarity 90.0%; Pred. No. 4.7e-34;  
 Matches 63; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSARGCGCQTSDNDAALAAASAAAALQIIREKIPEKYLFOFHNLSNL 60  
 DB 1 TLVWGEFQVDRSARGCGCQTSDNDAALAAASAAAALQIIREKIPEKYLFOFHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 drifdktppep 70  
 RESULT 10  
 AAB18678  
 ID AAB18678 standard; peptide; 70 AA.  
 XX AAB18678;  
 AC AAB18678;  
 XX 22-JAN-2001 (first entry)  
 DT Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX Synthetic.  
 OS Tomato golden mosaic virus.  
 XX Key Location/Qualifiers  
 FH Misc-difference 12  
 FT /note- "wild type residue replaced with Ala"  
 FT Misc-difference 13

FT Misc-difference 15  
 FT /note- "wild type residue replaced with Ala"  
 XX WO200054573-A1.  
 XX 21-SEP-2000.  
 XX 15-MAR-2000; 2000WO-US06759.  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT .  
 XX Claim 53; Page 42-43; 73pp; English.  
 PS The present sequence represents a mutant peptide, derived from a  
 XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX Sequence 70 AA:  
 SQ  
 Query Match 88.5%; Score 317; DB 21; Length 70;  
 Best Local Similarity 88.6%; Pred. No. 6.3e-34;  
 Matches 62; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSARGCGCQTSDNDAALAAASAAAALQIIREKIPEKYLFOFHNLSNL 60  
 DB 1 TLVWGEFQVDRSARGCGCQTSDNDAALAAASAAAALQIIREKIPEKYLFOFHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 DB 61 drifdktppep 70  
 RESULT 11  
 AAB18686  
 ID AAB18686 standard; peptide; 70 AA.  
 XX AAB18686;  
 AC AAB18686;  
 XX 22-JAN-2001 (first entry)  
 DT Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX Synthetic.  
 OS Tomato golden mosaic virus.  
 XX Key Location/Qualifiers  
 FH Misc-difference 12  
 FT /note- "wild type residue replaced with Ala"  
 FT Misc-difference 13









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:40:08 ; Search time 72.79 Seconds  
(without alignments)  
73.255 Million cell updates/sec

Title: US-09-289-346A-3  
Perfect score: 358  
Sequence: 1 TLVWGEFQVDRSGRGCCQT.....FQPHNLNSLDRIFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332	92.7	352	1 QOCVLI	ALI protein - toma
2	243	67.9	361	1 QOCVPT	ALI protein - pota
3	230	64.2	358	2 S07594	hypothetical prote
4	222	62.0	362	1 JQ1887	ALI protein - toma
5	217	60.6	355	1 QOCVW1	AV1 protein - abut
6	216	60.3	349	2 JQ2300	replicase - pepper
7	216	60.3	349	2 S31875	ALI protein - pepp
8	216	60.3	359	2 S39211	gene Cl protein -
9	214	59.8	358	1 JQ1870	ALI protein - toma
10	213	59.5	351	2 JQ2327	ALI protein - Indi
11	212	59.2	359	2 S22593	hypothetical prote
12	208	58.1	385	2 S28360	ALI protein - beet
13	207	57.8	360	2 S59885	replication-associ
14	203	56.7	359	2 S39235	gene Cl protein -
15	190	53.1	357	1 QOCVCI	ALI protein - toma
16	120	33.5	131	2 S45059	ALI protein (clone
17	115	32.1	347	1 QOCVSI	ALI protein - squa
18	69	19.3	577	2 S30237	transcription init
19	66	18.4	508	2 S59870	fork head domain p
20	66	18.4	840	2 T36175	probable large ATP
21	65	18.2	671	2 A35912	homeotic protein o
22	63	17.6	392	2 T44362	acetyl-CoA C-acety
23	62.5	17.5	700	2 T05841	spliceosome-associ
24	62.5	17.5	1077	2 A44067	serine-rich protei
25	62	17.3	409	2 T16781	hypothetical prote
26	61.5	17.2	327	2 F83408	hypothetical prote
27	61.5	17.2	705	2 T31157	hypothetical prote
28	61.5	17.2	1541	2 T02831	AAA protein I4171
29	61	17.0	339	2 T26328	hypothetical prote

30	61	17.0	371	2 B69502	alcohol dehydrogen
31	61	17.0	1265	2 T47626	structural main ten
32	61	17.0	1510	2 T33100	hypothetical prote
33	60.5	16.9	509	2 T02942	O-succinylhomoseri
34	60.5	16.9	829	2 A34692	ecdysone-induced p
35	60.5	16.9	883	2 S04722	puif 74E protein -
36	60	16.8	314	2 JCS273	paired type homeob
37	59.5	16.6	201	2 A81380	probable flagellar
38	59.5	16.6	244	2 B86197	hypothetical prote
39	59.5	16.6	470	2 T46814	gamma-aminobutyrat
40	59.5	16.6	481	2 A70091	probable phosphoes
41	59.5	16.6	509	2 T02940	O-succinylhomoseri
42	59	16.5	380	2 S70594	ubiquinol--cytochr
43	59	16.5	482	2 A39285	calcitonin recepto
44	59	16.5	498	2 I47130	calcitonin recepto
45	59	16.5	513	1 SYCHLE	5-aminolevulinate

## ALIGNMENTS

RESULT 1  
QOCVLI  
ALI protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Host: Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
C:Accession: A04170  
R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBO J. 3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <RAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 92.7%; Score 332; DB 1; Length 352;  
Best Local Similarity 92.9%; Pred. No. 7.2e-31;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCCQTSNDAAAALAAASAAAALQIIREKIPEKYLQFPHNLNSNL 60  
DB 111 TLVWGEFQVDRSGRGCCQTSNDAAAALAAASAAAALQIIREKIPEKYLQFPHNLNSNL 170  
QY 61 DRIFDKTPEP 70  
DB 171 DRIFDKTPEP 180

RESULT 2  
QOCVPT  
ALI protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: JU0364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: JU0362; MUID:91311403  
A:Accession: JU0364  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-361 <COU>  
A:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

Df 110 TLENGEFQIDGRSARGGQQSQANDAYAGALNTGSKSEALNVRLRELAPKDYLVLQFHNLNSNL 169

QY 61 DRI-----FDKTPPE 69  
||| ||| |

Df 170 DRITPPLEVYVSPFLSSDFRVEP 194

RESULT 5

QCVM1

A: Protein - abutilon mosaic virus (isolate West India)  
C: Species: abutilon mosaic virus  
C: Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994  
C: Accession: A36214  
R: Frischmuth, T.; Zimmatt, G.; Jeske, H.  
Virolgy 178, 461-468, 1990  
A: Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as  
A: Reference number: A36214; MUID: 91020984  
A: Accession: A36214  
A: Molecule type: DNA  
A: Residues: 1-355 <FRI>  
A: Cross-references: EMBL:X15983  
C: Geneticks:  
A: Map position: segment A  
C: Superfamily: tomato golden mosaic virus All protein

Query Match 60.6%; Score 217; DB 1; Length 355;  
Best Local Similarity 58.6%; Pred. No. 1.6e-17;  
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLWVGFEQVDGRSARGGCQTSDNAAAALAAASSAAAAAQIIRKEIPKYLQFPHNLNSNL 60  
| | | | | | | | | | | | | | | | | | | | : | : | | : | : | | |  
Df 110 TAENGFEQIDGRSARGGQQTANDSYAKALNAGDVQSALNLIKKEQPKDYVLQNHNIRSNA 169

QY 61 DRIFDKTPEP 70  
: | | | | | | |

Df 170 ERIFAKAPEP 179

RESULT 6

JQ2300

replicase - pepper huasteco virus (component A)  
N: Alternate names: ORF All protein  
C: Species: pepper huasteco virus  
C: Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 20-Sep-1999  
C: Accession: JQ2300  
R: Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,  
J. Gen. Virol. 74, 2225-2231, 1993  
A: Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b  
A: Reference number: JQ2299; MUID: 94015007  
A: Accession: JQ2300  
A: Molecule type: DNA  
A: Residues: 1-349 <TOR>  
A: Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856;1; PID:g61025  
C: Superfamily: tomato golden mosaic virus All protein

Query Match 60.3%; Score 216; DB 2; Length 349;  
Best Local Similarity 58.6%; Pred. No. 2e-17;  
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLWVGFEQVDGRSARGGCQTSDNAAAALAAASSAAAAAQIIRKEIPKYLQFPHNLNSNL 60  
| : | | | | | | | | | | | | | | | | | | | | : | : | | | | : | : | | |  
Df 110 TVWNGFEQIDGRSARGGQQSQANDTYAKALSASAEALQLIKKEQPHFLQFHNIVSNA 169

QY 61 DRIFDKTPEP 70  
: | | | | | | |

Df 170 NRIFQTPPEP 179

RESULT 7

S31875

All protein - pepper rizado amarillo virus



C:Species: pepper rizado amarillo virus  
C:Date: 22-Nov-1993 #sequence:revision 26-May-1995 #text\_change 20-Sep-1999  
C:Accession: S31875  
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Buonafina, P. et al. The complete nucleotide sequence of pepper huasteco virus: analysis and comparison with other members of the genus Tospovirus. *Journal of Virology* 1993  
A:Description: Complete nucleotide sequence of pepper huasteco virus: analysis and comparison with other members of the genus Tospovirus. *Journal of Virology* 1993  
A:Reference number: S31872

A:Accession: S31875  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-349 <TOR>  
A:Cross-references: EMBL:X70418; NID:G61023; PIDN:CAA49856.1; PID:G61025  
A:Note: the source is designated as pepper huasteco virus  
C:Superfamily: tomato golden mosaic virus ALL protein

```
Query Match      60.3%; Score 216; DB 2; Length 349;
Best Local Similarity 58.6%; Pred. No. 2e-17;
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;
```

**QY**

1 TLVWGEFOWDGRSARGCQTSDNAAAAAALASSAAAIQIIRKIKEVLFFQHNLNSL 60  
:- :::::||||||| :>::: || |::::: |: : ::::: ||  
**Dd** 110 TVEWGFEIDGRSARGGOASNDTYAKALSASAEEALOIKKEPOPHLOFINIVSNA 169

QY 61 DRFDKTPPEP 70  
: ||| |||  
DB 170 NRIFOTPPPEP 179

## RESULT

gene C1 protein - tomato yellow leaf curl virus  
 S39211  
 RESOLUT 0  
 C:Species: tomato yellow leaf curl virus  
 C:date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
 C:Accession: S39211  
 R:Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.  
 submitted to the EMBL Data Library, August 1993  
 A:Description: High similarity among the tomato yellow leaf curl virus isolates from the  
 A:Reference number: S39209

A: Reference number: S39209  
A: Accession: S39211  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-359 <NOR>  
C: Cross-references: EMBL:225751; NID:g433655; PIDN:CAA81026.1; PID:g433658  
C: Superfamily: tomato golden mosaic virus All protein

```
Query Match          60.3%; Score 216; DB 2; Length 359;
Best Local Similarity 58.0%; Pred. No. 2.1e-17;
Matches 40; Conservative 9; Mismatches 20; Indels 0; Gaps 0;
```

```
Qy 2 LVWGEFQVDRSARGCQTSNDAAAAAALASAAAAAQIIREKIPEKYLFFQHNLNSLD 61
    _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _
Db 111 LEWGTFTQIDGRSARGGQOTANDAYAKAINAGSKSEALVKEIAPROYTLHFHNINSND 170
```

Qy 62 RIFDKTPEP 70  
| : | | |  
Db 171 RVFQVPPAP 179

## RESULTS

JQ1870  
 Ali protein - tomato mottle virus (isolate Florida)  
 C:Species: tomato mottle virus  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
 C:Accession: JQ1870  
 R:Abouzid, A.M.; Polston, J.E.; Hiebert, E.  
 J. Gen. Virol. 73, 3225-3229, 1992  
 A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from  
 A:Reference number: JQ1869; MUID:J3107858  
 A:Accession: JQ1870  
 A:Status: translation not shown  
 A:Molecule type: DNA

A; Residues: 1-358 <ABO>  
A; Cross-references: GB:L14460  
C; Genetics:  
A; Map position: segment A  
C; Superfamily: tomato golden mosaic virus All protein

Query Match 59.8%; Score 214; DB 1; Length 358;  
Best Local Similarity 55.7%; Pred. NO. 3.5e-17;  
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

[illegible]

QY 61 DRIFDKTPEP 70  
: ||| | |||  
Db 167 ERIEAKAPEP 170

DEPT T 10

All protein - Indian cassava mosaic virus  
 N/Alternate names: replication-associated protein  
 C/Species: Indian cassava mosaic virus  
 C/Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
 C/Accession: JQ2327; S3583  
 R/Hong, Y.G.; Robinson, D.J.; Harrison, B.D.

J. Cell. Virol. 74, 2437-2443, 1993  
 A:Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted geminiviruses in the Americas  
 A:Reference number: JQ2326; PMID:94085670  
 A:Accession: JQ2327  
 A:Molecule type: DNA  
 A:Residues: 1-351 <HON>  
 A:Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046  
 C:Superfamily: tomato golden mosaic virus A11 protein

Query Match	59.5%	Score 213;	DB 2;	Length 351;
Best Local Similarity	61.2%;	Pred. No. 4.5e-17;		
Matches	41;	Conservative	10;	Mismatches 16;
				Indels 0;
				Gaps 0;

**QY**      4    WGEFQV DGRSARGGCOT\$NDAAAAAALAASSAAAALIIIRKIPKYLFQFHNLNSNLDRI    63  
        ||| :|||:||||| :|||·||| : | : |:|:| | + || |:|:|:|:|:|  
**pb**    113   WGTEIDGSRGGGOSANDAYAAAALNGSGSKSEALKILRELAPRDYLDRDFHISSNLDRI    172

QY	64	FDKTPPEP	70
Db	173	FTKPPPP	179

RESULT 11

R2391  
 R2392  
 R2393  
 R2394  
 R2395  
 R2396  
 R2397  
 R2398  
 R2399  
 R2400  
 R2401  
 R2402  
 R2403  
 R2404  
 R2405  
 R2406  
 R2407  
 R2408  
 R2409  
 R2410  
 R2411  
 R2412  
 R2413  
 R2414  
 R2415  
 R2416  
 R2417  
 R2418  
 R2419  
 R2420  
 R2421  
 R2422  
 R2423  
 R2424  
 R2425  
 R2426  
 R2427  
 R2428  
 R2429  
 R2430  
 R2431  
 R2432  
 R2433  
 R2434  
 R2435  
 R2436  
 R2437  
 R2438  
 R2439  
 R2440  
 R2441  
 R2442  
 R2443  
 R2444  
 R2445  
 R2446  
 R2447  
 R2448  
 R2449  
 R2450  
 R2451  
 R2452  
 R2453  
 R2454  
 R2455  
 R2456  
 R2457  
 R2458  
 R2459  
 R2460  
 R2461  
 R2462  
 R2463  
 R2464  
 R2465  
 R2466  
 R2467  
 R2468  
 R2469  
 R2470  
 R2471  
 R2472  
 R2473  
 R2474  
 R2475  
 R2476  
 R2477  
 R2478  
 R2479  
 R2480  
 R2481  
 R2482  
 R2483  
 R2484  
 R2485  
 R2486  
 R2487  
 R2488  
 R2489  
 R2490  
 R2491  
 R2492  
 R2493  
 R2494  
 R2495  
 R2496  
 R2497  
 R2498  
 R2499  
 R2500  
 R2501  
 R2502  
 R2503  
 R2504  
 R2505  
 R2506  
 R2507  
 R2508  
 R2509  
 R2510  
 R2511  
 R2512  
 R2513  
 R2514  
 R2515  
 R2516  
 R2517  
 R2518  
 R2519  
 R2520  
 R2521  
 R2522  
 R2523  
 R2524  
 R2525  
 R2526  
 R2527  
 R2528  
 R2529  
 R2530  
 R2531  
 R2532  
 R2533  
 R2534  
 R2535  
 R2536  
 R2537  
 R2538  
 R2539  
 R2540  
 R2541  
 R2542  
 R2543  
 R2544  
 R2545  
 R2546  
 R2547  
 R2548  
 R2549  
 R2550  
 R2551  
 R2552  
 R2553  
 R2554  
 R2555  
 R2556  
 R2557  
 R2558  
 R2559  
 R2560  
 R2561  
 R2562  
 R2563  
 R2564  
 R2565  
 R2566  
 R2567  
 R2568  
 R2569  
 R2570  
 R2571  
 R2572  
 R2573  
 R2574  
 R2575  
 R2576  
 R2577  
 R2578  
 R2579  
 R2580  
 R2581  
 R2582  
 R2583  
 R2584  
 R2585  
 R2586  
 R2587  
 R2588  
 R2589  
 R2590  
 R2591  
 R2592  
 R2593  
 R2594  
 R2595  
 R2596  
 R2597  
 R2598  
 R2599  
 R2600  
 R2601  
 R2602  
 R2603  
 R2604  
 R2605  
 R2606  
 R2607  
 R2608  
 R2609  
 R2610  
 R2611  
 R2612  
 R2613  
 R2614  
 R2615  
 R2616  
 R2617  
 R2618  
 R2619  
 R2620  
 R2621  
 R2622  
 R2623  
 R2624  
 R2625  
 R2626  
 R2627  
 R2628  
 R2629  
 R2630  
 R2631  
 R2632  
 R2633  
 R2634  
 R2635  
 R2636  
 R2637  
 R2638  
 R2639  
 R2640  
 R2641  
 R2642  
 R2643  
 R2644  
 R2645  
 R2646  
 R2647  
 R2648  
 R2649  
 R2650  
 R2651  
 R2652  
 R2653  
 R2654  
 R2655  
 R2656  
 R2657  
 R2658  
 R2659  
 R2660  
 R2661  
 R2662  
 R2663  
 R2664  
 R2665  
 R2666  
 R2667  
 R2668  
 R2669  
 R2670  
 R2671  
 R2672  
 R2673  
 R2674  
 R2675  
 R2676  
 R2677  
 R2678  
 R2679  
 R2680  
 R2681  
 R2682  
 R2683  
 R2684  
 R2685  
 R2686  
 R2687  
 R2688  
 R2689  
 R2690  
 R2691  
 R2692  
 R2693  
 R2694  
 R2695  
 R2696  
 R2697  
 R2698  
 R2699  
 R2700  
 R2701  
 R2702  
 R2703  
 R2704  
 R2705  
 R2706  
 R2707  
 R2708  
 R2709  
 R2710  
 R2711  
 R2712  
 R2713  
 R2714  
 R2715  
 R2716  
 R2717  
 R2718  
 R2719  
 R2720  
 R2721  
 R2722  
 R2723  
 R2724  
 R2725  
 R2726  
 R2727  
 R2728  
 R2729  
 R2730  
 R2731  
 R2732  
 R2733  
 R2734  
 R2735  
 R2736  
 R2737  
 R2738  
 R2739  
 R2740  
 R2741  
 R2742  
 R2743  
 R2744  
 R2745  
 R2746  
 R2747  
 R2748  
 R2749  
 R2750  
 R2751  
 R2752  
 R2753  
 R2754  
 R2755  
 R2756  
 R2757  
 R2758  
 R2759  
 R2760  
 R2761  
 R2762  
 R2763  
 R2764  
 R2765  
 R2766  
 R2767  
 R2768  
 R2769  
 R2770  
 R2771  
 R2772  
 R2773  
 R2774  
 R2775  
 R2776  
 R2777  
 R2778  
 R2779  
 R2780  
 R2781  
 R2782  
 R2783  
 R2784  
 R2785  
 R2786  
 R2787  
 R2788  
 R2789  
 R2790  
 R2791  
 R2792  
 R2793  
 R2794  
 R2795  
 R2796  
 R2797  
 R2798  
 R2799  
 R2800  
 R2801  
 R2802  
 R2803  
 R2804  
 R2805  
 R2806  
 R2807  
 R2808  
 R2809  
 R2810  
 R2811  
 R2812  
 R2813  
 R2814  
 R2815  
 R2816  
 R2817  
 R2818  
 R2819  
 R2820  
 R2821  
 R2822  
 R2823  
 R2824  
 R2825  
 R2826  
 R2827  
 R2828  
 R2829  
 R2830  
 R2831  
 R2832  
 R2833  
 R2834  
 R2835  
 R2836  
 R2837  
 R2838  
 R2839  
 R2840  
 R2841  
 R2842  
 R2843  
 R2844  
 R2845

A:Accession: 542355  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-359 <KHE>  
A:Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match	59.2%	Score 212;	DB 2;	Length 359;
Best Local Similarity	56.5%;	Pred. No. 6e-17;		
Matches	39;	Conservative	10;	Mismatches 20;
				Indels 0;
				Gaps 0;





---

**THIS PAGE BLANK (USPTO)**

.





```

CC      -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X15983; -: NOT_ANNOTATED_CDS.
DR      PIR; A36214; OOCVWL
DR      InterPro; IPR001191; Gemini_AL1.
DR      Pfam; PF00799; Gemini_AL1; 1.
DR      PRINTS; PR00227; GEMCOATALL1.
DR      PRINTS; PR00228; GEMCOATCLVL1.
DR      ProDom; PD000736; Gemini_AL1; 1.
KW      ATP-binding. 221 228  ATP (POTENTIAL).
FT      NP_BIND
SQ      SEQUENCE 355 AA: 40257 MW; 16A2CAB63251E95 CRC64;

Query Match      60.6%; Score 217; DB 1; Length 355;
Best Local Similarity 58.6%; Pred No. 8.3e-18;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY      1 TLVWGEFQVDSARGGCGTSDNDAALAAASSAAALQIIREKIPEKYLQFHNLSNL 60
      ++++++
DB      110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSLNLKEQPKDYVLQFHNLSNL 169

QY      61 DRIFDKTPEP 70
      ++++++
DB      170 ERIFAKAPEP 179

RESULT 7
VAL1_PHVU
ID      VAL1_PHVU STANDARD; PRT; 349 AA.
AC      Q06923;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      01-JUN-1994 (Rel. 29, Last annotation update)
DE      AL1 PROTEIN.
GN      AL1.
OS      Pepper huasteco virus (PHV).
OC      Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX      NCBI_TaxID=28349;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94015007; PubMed=8409944;
RA      Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA      Rivera-Bustamante R.F.;
RT      "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT      comparison with bipartite geminiviruses.";
RL      J. Gen. Virol. 74:2225-2231(1993).
CC      -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X70418; CAA49856.1; -
DR      PIR; S31875; S31875.
DR      PIR; JQ2300; JQ2300.
DR      InterPro; IPR001191; Gemini_AL1.
DR      Pfam; PF00799; Gemini_AL1; 1.
DR      PRINTS; PR00227; GEMCOATALL1.
DR      PRINTS; PR00228; GEMCOATCLVL1.
DR      ProDom; PD000736; Gemini_AL1; 1.

QY      1 TLVWGEFQVDSARGGCGTSDNDAALAAASSAAALQIIREKIPEKYLQFHNLSNL 60
      ++++++
DB      110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSLNLKEQPKDYVLQFHNLSNL 169

QY      61 DRI-----FDKTP 69
      ++++++
DB      170 DRIFTPPLEVYVPSFLSSSFDKTP 174

RESULT 6
VAL1_ABMW
ID      VAL1_ABMW STANDARD; PRT; 355 AA.
AC      P21947;
DT      01-AUG-1991 (Rel. 19, Created)
DT      01-AUG-1991 (Rel. 19, Last sequence update)
DT      01-JUN-1994 (Rel. 29, Last annotation update)
DE      AL1 PROTEIN.
GN      AC1.
OS      Abutilon mosaic virus (isolate West India).
OC      Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX      NCBI_TaxID=10816;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91020984; PubMed=2219703;
RA      Frischmuth T., Zimmatt G., Joske H.;
RT      "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT      as well as eukaryotic features.";
RL      Virology 178:461-468(1990).

```

```

CC      QY      1 TLVWGEFQVDGSRARGCOTSNDAALAAASSAAAALQIIREKIPEKYLFQFHNLNSL 60
CC      DB      109 TVEWGFQIDGRSARGQGSANDAYAKALNCSKSEALNVRELVPKDVLYQFHNLNSL 168
CC      QY      61 DRIFDKTPEP 70
CC      DB      169 DRIFQEPPAP 178
CC
CC      RESULT 5
CC      VAL1_TYLCA STANDARD; PRT; 362 AA.
CC      AC      P36279;
CC      DT      01-JUN-1994 (Rel. 29, Created)
CC      DT      01-AUG-1991 (Rel. 29, Last sequence update)
CC      DT      20-AUG-2001 (Rel. 40, Last annotation update)
CC      DE      AL1 PROTEIN (C1 PROTEIN).
CC      GN      C1.
CC      OS      Tomato yellow leaf curl virus (strain Australia) (TYLCV).
CC      OC      Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
CC      OX      NCBI_TaxID=36447;
CC      RN      [1]
CC      RP      SEQUENCE FROM N.A.
CC      RX      MEDLINE=93139778; PubMed=8423446;
CC      RT      Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
CC      RT      "Nucleotide sequence and genome organization of tomato leaf curl
CC      RT      geminivirus.";
CC      RL      J. Gen. Virol. 74:147-151(1993).
CC      CC      -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC      DR      InterPro: JQ1887; JQ1887;
CC      DR      InterPro: IPR001191; Gemini_AL1.
CC      DR      Pfam: PF00799; Gemini_AL1; 1.
CC      DR      PRINTS; PR00227; GEMCOATALL.
CC      DR      PRINTS; PR00228; GEMCOATCLVL1.
CC      DR      ProDom: PD000736; GEMINI_AL1; 1.
CC      KW      ATP-binding. 221 228
CC      FT      NP_BIND
CC      SQ      SEQUENCE 362 AA; 41197 MW; 343E7184B4704098 CRC64;

Query Match          62.0%; Score 222; DB 1; Length 362;
Best Local Similarity 52.9%; Pred.No. 2.2e-18;
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

CC      QY      1 TLVWGEFQVDGSRARGCOTSNDAALAAASSAAAALQIIREKIPEKYLFQFHNLNSL 60
CC      DB      110 TLEWGEFQIDGRSARGQGSANDAYAKALNTCKSEALNVRELAPKDYVLYQFHNLNSL 169
CC      QY      61 DRI-----FDKTPE 69
CC      DB      170 DRITPPLEVYVSFLSSSFDRVPE 194
CC
CC      RESULT 6
CC      VAL1_ABMVW STANDARD; PRT; 355 AA.
CC      AC      P21947;
CC      DT      01-AUG-1991 (Rel. 19, Created)
CC      DT      01-AUG-1991 (Rel. 19, Last sequence update)
CC      DT      01-JUN-1994 (Rel. 29, Last annotation update)
CC      DE      AL1 PROTEIN.
CC      GN      AC1.
CC      OS      Abutilon mosaic virus (isolate West India).
CC      OC      Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
CC      OX      NCBI_TaxID=10816;
CC      RN      [1]
CC      RP      SEQUENCE FROM N.A.
CC      RX      MEDLINE=91020984; PubMed=2219703;
CC      RT      Frischmuth T., Zimmatt G., Joske H.;
CC      RT      "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
CC      RT      as well as eukaryotic features.";
CC      RL      Virology 178:461-468(1990).

```





CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L14460; AAC32414.1; -;  
DR PIR; J01870; JQ1870.  
DR InterPro; IPR001191; Gemini\_ALL1.  
DR Pfam; PF00799; Gemini\_ALL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLL.  
DR ProDom; PD000736; Gemini\_ALL1; 1.  
KW ATP-binding.  
FT NP\_BIND 222 229 ATP (BY SIMILARITY).  
SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEEAC6950 CRC64;

Query Match 59.8%; Score 214; DB 1; Length 361;  
Best Local Similarity 55.7%; Pred. No. 1.9e-17;  
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLWGEFQVDRSARGCGTSDNDAALAAASSAAALQIIREKIPEKYLQFHHNLSNL 60  
Db 110 TIEWGDFQIDGRSARGGQSANDSYAKALNAGSVQSLAVLREEQPKDFVLQNHNRSLN 169  
Qy 61 DRFDKTPPEP 70  
Db 170 ERIFAKAPEP 179

## RESULT 11

VAL1\_TYLCM STANDARD; PRT; 359 AA.  
AC P27260;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE ALL PROTEIN (C1 PROTEIN).  
GN C1.  
OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92107660; PubMed=1840676;  
RA Kheyr-Pour A., Bandahmane M., Matzelt V., Accotto G.P., Crespi S.,  
RA Gronenborn B.  
RT "Tomato yellow leaf curl virus from Sardinia is a  
RT whitefly-transmitted monopartite geminivirus."  
RL Nucleic Acids Res. 19:6763-6769(1991).

CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X61153; CA443466.1; -;  
DR PIR; S22593; S22593.  
DR InterPro; IPR001191; Gemini\_ALL1.  
DR Pfam; PF00799; Gemini\_ALL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLL.  
DR ProDom; PD000736; Gemini\_ALL1; 1.  
KW ATP-binding.  
FT NP\_BIND 220 227 ATP (POTENTIAL).  
SQ SEQUENCE 359 AA; 40733 MW; 9717B4A07C93EFA7 CRC64;

Query Match 59.2%; Score 212; DB 1; Length 359;  
Best Local Similarity 56.5%; Pred. No. 3.2e-17;  
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 2 LVMGEFQVDRSARGCGTSDNDAALAAASSAAALQIIREKIPEKYLQFHHNLSNL 61  
Db 111 LEWGTFFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVYLHFNHNSL 170  
Qy 62 RIFDKTPPEP 70  
Db 171 KVFOVPPAP 179

## RESULT 12

VAL1\_BCTV STANDARD; PRT; 358 AA.  
ID VAL1\_BCTV  
AC P14991;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE ALL PROTEIN (40.8 KDA PROTEIN).  
OS Beet curly top virus (BCTV).  
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.  
OX NCBI\_TaxID=10840;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;  
RT "The nucleotide sequence of an infectious clone of the geminivirus  
RT beet curly top virus."  
RL EMBO J. 5:1761-1767(1986).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL1 PROTEIN FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X04144; -; NOT ANNOTATED CDS.  
DR InterPro; IPR001191; Gemini\_ALL1.  
DR Pfam; PF00799; Gemini\_ALL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLL.  
DR ProDom; PD000736; Gemini\_ALL1; 1.  
KW ATP-binding.  
FT NP\_BIND 222 229 ATP (POTENTIAL).  
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 58.1%; Score 208; DB 1; Length 358;  
Best Local Similarity 55.7%; Pred. No. 9.1e-17;  
Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

Qy 1 TLWGEFQVDRSARGCGTSDNDAALAAASSAAALQIIREKIPEKYLQFHHNLSNL 60  
Db 110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEQPKDYFLQHHNLLNA 169  
Qy 61 DRFDKTPPEP 70  
Db 170 QKIFORPPDP 179

## RESULT 13

VAL1\_TYLCV STANDARD; PRT; 357 AA.  
ID VAL1\_TYLCV  
AC P27259;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)

```
DE  ALL PROTEIN (C1 PROTEIN).
GN  C1.
OS  Tomato yellow leaf curl virus (TYLCV).
OC  Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
NCBI_TaxID=10832;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92024070; PubMed=1926771;
RA  Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT  "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
with a single genomic component.";
RL  Virology 185:151-161(1991).
CC  -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X15656; CAA33688.1; .
DR  PIR; D40779; QOCVCL.
DR  InterPro; IPR001191; Gemini_ALI.
DR  Pfam; PF00799; Gemini_ALI; 1.
DR  PRINTS; PR00227; GEMCOATALL.
DR  PRINTS; PR00228; GEMCOATCLVLL.
DR  ProDom; PD000736; Gemini_ALI; 1.
KW  ATP-binding. 219 226 ATP (POTENTIAL).
FT  NP_BIND 219 226
SQ  SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;

Query Match 53.1%; Score 190; DB 1; Length 357;
Best Local Similarity 60.7%; Pred. No. 1.le-14;
Matches 37; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY  4 WGFQVDGRSARGCGQTSNDAAAALAAASSAAALQIIREKIPEKYLFQFHNLSNLDRI 63
DB  111 FGVSDIGRSARGCGQSANDAYAEALNSGSKSEALNLEKAPKDYILQFHNLSNLDRI 170

QY  64 F 64
DB  171 F 171

RESULT 14
VALI_SLCV STANDARD; PRT; 347 AA.
AC  P29048;
DT  01-DEC-1992 (Rel. 24, Created)
DT  01-DEC-1992 (Rel. 24, Last sequence update)
DT  01-DEC-1992 (Rel. 24, Last annotation update)
DE  ALL PROTEIN.
OS  Squash leaf curl virus.
OC  Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX  NCBI_TaxID=10829;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=91082449; PubMed=1984668;
RA  Lazarowitz S.G., Iazdins I.B.;
RT  "Infectivity and complete nucleotide sequence of the cloned genomic
components of a bipartite squash leaf curl geminivirus with a broad
host range phenotype.";
RL  Virology 180:58-69(1991).
CC  -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
```

```
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M38183; AAC32410.1; ALT_INIT.
DR  PIR; C36785; QOCV51.
DR  InterPro; IPR001191; Gemini_ALI.
DR  Pfam; PF00799; Gemini_ALI; 1.
DR  PRINTS; PR00227; GEMCOATALL.
DR  PRINTS; PR00228; GEMCOATCLVLL.
DR  ProDom; PD000736; Gemini_ALI; 1.
KW  ATP-binding. 218 225 ATP (POTENTIAL).
FT  NP_BIND 218 225
SQ  SEQUENCE 347 AA; 39110 MW; AFDAEBDE122110E CRC64;

Query Match 32.1%; Score 115; DB 1; Length 347;
Best Local Similarity 36.4%; Pred. No. 5e-06;
Matches 24; Conservative 14; Mismatches 24; Indels 4; Gaps 1;

QY  5 GEFQVDGRSARGCGQTSNDAAAALAAASSAAALQIIREKIPEKYLFQFHNLSNLDRI 64
DB  111 GQYKVS-----GSKSNKDDVYHNAVNSAGSAGEALDIKAGDPKTFIVYHNLNLANVERLF 171

QY  65 DKTPEP 70
DB  172 QKPPEP 177

RESULT 15
T2FA_DROME STANDARD; PRT; 577 AA.
ID  T2FA_DROME
AC  Q05913; O9V119;
DT  01-FEB-1994 (Rel. 28, Created)
DT  20-AUG-2001 (Rel. 40, Last sequence update)
DT  20-AUG-2001 (Rel. 40, Last annotation update)
DE  TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT (TFIIF-ALPHA)
DE  (TRANSCRIPTION INITIATION FACTOR 5, LARGE CHAIN) (TF5A).
GN  TFIIF-ALPHA OR CG10281.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93219133; PubMed=8464745;
RA  Gong D.-W., Horikoshi M., Nakatani Y.;
RT  "Analysis of cDNA encoding Drosophila transcription initiation factor
TFIIF alpha (RAP74).";
RL  Nucleic Acids Res. 21:1492-1492(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93219101; PubMed=8464716;
RA  Kephart D., Price M.P., Burton Z.F., Finkelstein A.,
RA  Greenblatt J., Price D.H.;
RT  "Cloning of a Drosophila cDNA with sequence similarity to human
transcription factor RAP74.";
RL  Nucleic Acids Res. 21:1319-1319(1993).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BERKELEY;
RX  MEDLINE=20196006; PubMed=10731132;
RA  Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins K.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA  Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
```

Search completed: January 3, 2002, 15:57:15  
Job time: 1101 sec

```

Query Match      19.3%; Score 69; DB 1; Length 577;
Best Local Similarity 36.4%; Pred. No. 1.8;
Matches 20; Conservative 9; Mismatches 20; Indels 6; Gaps 1;

QY 13 SARGCGCOTSDAAALAAASAAALQIIRE-----KPEKYLQFPHNLSNLD 61
  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 10 SAAGSGTSAIAAAASAAASVAGSSANNOVEFKTRVPMKPKHHVMRFNATLND 64
  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:56:24 ; Search time 131.69 seconds  
(without alignments)  
77.751 Million cell updates/sec

Title: US-09-289-346A-3

Perfect score: 358

Sequence: 1 TLVWGEFQVDRSGRGCCQT.....FQFHNLNSLDRIFDKTPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.todent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	291	81.3	352	12 Q9E000	Q9E000 tomato rugo
2	276	77.1	226	12 Q09727	Q09727 leonurus mo
3	271	75.7	314	12 Q9ELT8	Q9ELT8 sweet potat
4	271	75.7	364	12 Q9QS55	Q9QS55 sweet potat
5	264	73.7	226	12 Q9WHF6	Q9WHF6 tomato mild
6	262	73.2	361	12 Q87574	Q87574 bean golden
7	260	72.6	149	12 P88975	P88975 macroptiliu
8	260	72.6	225	12 Q9QDB1	Q9QDB1 cowpea gold
9	260	72.6	233	12 Q9YLA4	Q9YLA4 macroptiliu
10	259	72.3	185	12 Q98693	Q98693 sida golden
11	251	70.1	190	12 Q9Z089	Q9Z089 tobacco lea
12	251	70.1	190	12 Q9Z084	Q9Z084 tobacco lea
13	250	69.8	190	12 Q9W827	Q9W827 tobacco lea
14	250	69.8	208	12 Q9Z0C4	Q9Z0C4 tobacco lea
15	250	69.8	208	12 Q9Z0C0	Q9Z0C0 tobacco lea
16	250	69.8	208	12 Q9Z0B8	Q9Z0B8 tobacco lea
17	249	69.6	208	12 Q9Z0A0	Q9Z0A0 tobacco lea
18	248	69.3	208	12 Q9Z0C6	Q9Z0C6 tobacco lea
19	247	69.0	362	12 Q56816	Q56816 chayote mos

20	246	68.7	208	12 Q9Z0B6	Q9Z0B6 tobacco lea
21	245	68.4	203	12 Q9Z083	Q9Z083 tobacco lea
22	245	68.4	363	12 Q73577	Q73577 cotton leaf
23	244	68.2	362	12 Q9IN48	Q9IN48 okra enatio
24	243	67.9	234	12 Q39180	Q39180 potato yell
25	242	67.6	231	12 Q96620	Q96620 african tom
26	242	67.6	353	12 Q72692	Q72692 beet curly
27	242	67.6	354	12 Q65438	Q65438 beet curly
28	242	67.6	359	12 Q9Y2V4	Q9Y2V4 tomato yell
29	242	67.6	359	12 Q9Y2V2	Q9Y2V2 tomato yell
30	242	67.6	359	12 Q9YUX7	Q9YUX7 tomato yell
31	242	67.6	359	12 Q9YL27	Q9YL27 tomato yell
32	241	67.3	358	12 Q65418	Q65418 beet curly
33	241	67.3	359	12 Q88942	Q88942 tomato yell
34	240	67.0	360	12 Q9DX10	Q9DX10 ageratatum ve
35	239	66.8	307	12 Q91ET7	Q91ET7 cotton leaf
36	239	66.8	361	12 Q72723	Q72723 cotton leaf
37	238	66.5	190	12 Q9Z0A7	Q9Z0A7 tobacco lea
38	238	66.5	231	12 Q67609	Q67609 tomato leaf
39	237	66.2	190	12 Q9Z086	Q9Z086 tobacco lea
40	234	65.4	363	12 Q72705	Q72705 cotton leaf
41	234	65.4	363	12 Q72719	Q72719 cotton leaf
42	233	65.1	362	12 Q9YVT7	Q9YVT7 althea rose
43	233	65.1	363	12 Q73494	Q73494 okra yellow
44	232	64.8	361	12 Q99DR1	Q99DR1 chilli leaf
45	231	64.5	306	12 Q91ET1	Q91ET1 cotton leaf

## ALIGNMENTS

RESULT 1  
Q9E000 PRELIMINARY; PRT; 352 AA.  
ID Q9E000;  
AC Q9E000;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE REPLICATION-ASSOCIATED PROTEIN.  
GN AC1  
OS Tomato rugose mosaic virus.  
OC Viruses; SSNDA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=134599;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,  
Zambolim E.M., Zerbini F.M.;  
RT "Molecular Cloning and Characterization of Tomato rugose mosaic virus  
(TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,  
RT Minas Gerais, Brazil."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF291705; AAC15546.1;  
DR InterPro; IPR001191; Geminivir\_1.  
DR Pfam; PF00799; Geminivir\_1.  
DR PRINTS; PR00227; GEMCONTALL.  
DR ProDom; PD000736; Geminivir\_1.  
SQ SEQUENCE 352 AA; 40012 MW; 47CD55838E24D613 CRC64;

Query Match 81.3%; Score 291; DB 12; Length 352;  
Best Local Similarity 78.6%; Pred. No. 1.6e-25;  
Matches 55; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCCQTSNDAAAALAAASSAAAAALQIIREKIPEKYLFOFHNLNSL 60  
DB 111 TLVWGEFQIDRGSGCCQTANDAAAALNAPSKDIALQIIRKMPKFLFQFHNLNSL 170  
QY 61 DRIFDKTPEP 70  
DB 171 DRIFARAPEP 180

RESULT 2

```
009727
ID Q09727 PRELIMINARY; PRT; 226 AA.
AC Q09727;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaisc virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 77.1%; Score 276; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 5.3e-24;
Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAALQIIREKPEKYLQFPHNLSNL 60
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111 TVMEGEFQVDRSARGCGQTVNDAAAFALNPKRYALQIIREKPEKYLQFPHNLSNL 170

QY 61 DRIFDKTP 70
   ||||| |||||
Db 171 DRIFAKAPEP 180

RESULT 3
Q9ELT8 PRELIMINARY; PRT; 314 AA.
ID Q9ELT8;
AC Q9ELT8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATION PROTEIN.
GN AC1.
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=100755;
RN [1]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT "Detection of a geminivirus infecting sweet potato in the United States.";
RL Plant Dis. 82:1253-1257(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF288227; AAG01006.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
SQ SEQUENCE 314 AA; 35153 MW; 6BG220613046943F CRC64;

Query Match 75.7%; Score 271; DB 12; Length 314;
Best Local Similarity 79.4%; Pred. No. 2.9e-23;
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAALQIIREKPEKYLQFPHNLSNL 60
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 TITWGEFQVDRSARGCGQQTANDAAALNAGSKEAALQIIREKPEKYLQFPHNLSNL 169

QY 61 DRIFDKTP 70
   ||||| |||||
Db 170 DRIFSPPP 177

RESULT 5
Q9WHF6 PRELIMINARY; PRT; 226 AA.
ID Q9WHF6;
AC Q9WHF6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN96-H5KW;
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131071; AAD33471.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
FT NON_TER 226
```

```
Db 110 TITWGEFQVDRSARGCGQQTANDAAALNAGSKEAALQIIREKPEKYLQFPHNLSNL 169
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 DRIFDKTP 68
   ||||| |||||
Db 170 DRIFSPPP 177

RESULT 4
Q9QS55 PRELIMINARY; PRT; 364 AA.
ID Q9QS55;
AC Q9QS55;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION INITIATION PROTEIN AC1.
GN AC1.
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=100755;
RN [1]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT "Detection of a geminivirus infecting sweet potato in the United States.";
RL Plant Dis. 82:1253-1257(1998).
DR EMBL; AF104036; AAD47173.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 75.7%; Score 271; DB 12; Length 364;
Best Local Similarity 79.4%; Pred. No. 3.4e-23;
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAALQIIREKPEKYLQFPHNLSNL 60
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 TITWGEFQVDRSARGCGQQTANDAAALNAGSKEAALQIIREKPEKYLQFPHNLSNL 169

QY 61 DRIFDKTP 68
   ||||| |||||
Db 170 DRIFSPPP 177

RESULT 5
Q9WHF6 PRELIMINARY; PRT; 226 AA.
ID Q9WHF6;
AC Q9WHF6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN96-H5KW;
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131071; AAD33471.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
FT NON_TER 226
```

SQ SEQUENCE 226 AA; 25941 MW; 2EA41116712871A23 CRC64;  
 Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAMAICAN;  
 RA ROYE M.E.;  
 RL Thesis (1996), Biochemistry, University of the West Indies, Jamaica.  
 DR EMBL: U75278; AAB36919.1; -;  
 DR InterPro: IPR001191; Gemin\_A1.  
 DR Pfam: PF00799; Gemin\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin\_A1; 1.  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 16785 MW; E4CF5BED4C9CD508 CRC64;  
 Query Match 72.6%; Score 260; DB 12; Length 149;  
 Best Local Similarity 68.6%; Pred. No. 2.3e-22;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 TLWGFQVQDGRSARGGCGTSDNDAAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 60  
 DB 111 TIEWGFQIDGRSARGGCGTSDNDAAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 170  
 QY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180  
 RESULT 6  
 ID Q67574 PRELIMINARY; PRT; 361 AA.  
 AC Q67574;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN All.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 Maxwell D.P., Russell D.R.;  
 RL Phytopathology 81:980-985(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 Morales F.J., Maxwell D.P.;  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M88686; AA46312.1; -;  
 DR InterPro: IPR001191; Gemin\_A1.  
 DR Pfam: PF00799; Gemin\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin\_A1; 1.  
 SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;  
 Query Match 73.2%; Score 262; DB 12; Length 361;  
 Best Local Similarity 73.1%; Pred. No. 3.6e-22;  
 Matches 49; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
 QY 4 WGFQVQDGRSARGGCGTSDNDAAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 63  
 DB 113 WGFQVQDGRSARGGCGTSDNDAAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 172  
 QY 64 FDKTPEP 70  
 DB 173 FTRAPDP 179  
 RESULT 7  
 ID P88975 PRELIMINARY; PRT; 149 AA.  
 AC P88975;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN AC1.  
 OS Macrotidium golden mosaic geminivirus.  
 OS

SQ SEQUENCE 226 AA; 25941 MW; 2EA41116712871A23 CRC64;  
 Viruses: ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAMAICAN;  
 RA ROYE M.E.;  
 RL Thesis (1996), Biochemistry, University of the West Indies, Jamaica.  
 DR EMBL: U75278; AAB36919.1; -;  
 DR InterPro: IPR001191; Gemin\_A1.  
 DR Pfam: PF00799; Gemin\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin\_A1; 1.  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 16785 MW; E4CF5BED4C9CD508 CRC64;  
 Query Match 72.6%; Score 260; DB 12; Length 149;  
 Best Local Similarity 68.6%; Pred. No. 2.3e-22;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 TLWGFQVDRSARGGCGTSDNDAAAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 60  
 DB 111 TIEWGFQIDGRSARGGQGTANDAAAEALNASSKEAMRIKEKLPKFLQYHNLSSNL 170  
 QY 61 DRIFDKTPEP 70  
 DB 171 DRIFAKAPEP 180  
 RESULT 6  
 ID Q67574 PRELIMINARY; PRT; 361 AA.  
 AC Q67574;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE PUTATIVE REPLICATIVE PROTEIN.  
 GN All.  
 OS Bean golden mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 NCBI\_TaxID=10839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,  
 Maxwell D.P., Russell D.R.;  
 RL Phytopathology 81:980-985(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,  
 Morales F.J., Maxwell D.P.;  
 RL Plant Dis. 75:336-342(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M88686; AA46312.1; -;  
 DR InterPro: IPR001191; Gemin\_A1.  
 DR Pfam: PF00799; Gemin\_A1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemin\_A1; 1.  
 SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;  
 Query Match 73.2%; Score 262; DB 12; Length 361;  
 Best Local Similarity 73.1%; Pred. No. 3.6e-22;  
 Matches 49; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
 QY 4 WGFQVDRSARGGCGTSDNDAAAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 63  
 DB 113 WGFQVDRSARGGQGTANDAAAEALNASSKEAMRIKEKLPKFLQYHNLSSNL 172  
 QY 64 FDKTPEP 70  
 DB 173 FTRAPDP 179  
 RESULT 7  
 ID P88975 PRELIMINARY; PRT; 149 AA.  
 AC P88975;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN AC1.  
 OS Macrotidium golden mosaic geminivirus.  
 OS





```
RL J. Plant Res. 110:247-257(1997).
DR EMBL; AB001318; BAA34039.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRODOM; PD000736; Gemini_AL1; 1.
FT NON_TER 1
FT NON_TER 190
SQ SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

Query Match 70.1%; Score 251; DB 12; Length 190;
Best Local Similarity 58.8%; Pred. No. 3.3e-21;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDRSARGGCOTSNDAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 60
Db 85 TLEWGTFOIDGRSARGGCQNADCAEALNAASKAEALSIIREKLPKDFIFOYHNLSNL 144

QY 61 DRI-----FDKTP 69
Db 145 DRIFAPPLEVFCPTASSFDQVPE 169

RESULT 13
QY 9W827 PRELIMINARY; PRT; 190 AA.
AC QW827;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE C1 PROTEIN (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RC STRAIN=GORAL;
SEQUENCE FROM N.A.
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL; AB001303; BAA34010.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRODOM; PD000736; Gemini_AL1; 1.
FT NON_TER 1
FT NON_TER 190
SQ SEQUENCE 190 AA; 21444 MW; 93C3742A8EBDB7EB CRC64;

Query Match 69.8%; Score 250; DB 12; Length 190;
Best Local Similarity 58.8%; Pred. No. 4.3e-21;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDRSARGGCOTSNDAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 60
Db 85 TLEWGTFOIDGRSARGGCQNADCAEALNAASKAEALSIIREKLPKDFIFOYHNLSNL 144

QY 61 DRI-----FDKTP 69
Db 145 DRIFAPPLEVFCPTASSFDQVPE 169

RESULT 14
QY 9Z0C4 PRELIMINARY; PRT; 208 AA.
AC Q9Z0C4;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL AND COMPLETE CDS (FRAGMENT).
```

```
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ABURA3;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL; AB001294; BAA33992.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRODOM; PD000736; Gemini_AL1; 1.
FT NON_TER 1
FT NON_TER 208
SQ SEQUENCE 208 AA; 23526 MW; 249CC31D8729C72D CRC64;

Query Match 69.8%; Score 250; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 4.8e-21;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDRSARGGCOTSNDAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 60
Db 102 TLEWGTFOIDGRSARGGCQNADCAEALNAASKAEALSIIREKLPKDFIFOYHNLSNL 161

QY 61 DRI-----FDKTP 69
Db 162 DRIFAPPLEVFCPTASSFDQVPE 186

RESULT 15
QY 9Z0C0 PRELIMINARY; PRT; 208 AA.
AC Q9Z0C0;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE C1 AND C4 GENES, CLONE AMG-1(B152), PARTIAL AND COMPLETE CDS (FRAGMENT).
GN C1.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AMGB152;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL J. Plant Res. 110:247-257(1997).
DR EMBL; AB001296; BAA33996.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRODOM; PD000736; Gemini_AL1; 1.
FT NON_TER 1
FT NON_TER 208
SQ SEQUENCE 208 AA; 23486 MW; E301135F799C3DAD CRC64;

Query Match 69.8%; Score 250; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 4.8e-21;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDRSARGGCOTSNDAALAAASSAAALQIIREKIPEKYLFOFHNLSNL 60
Db 102 TLEWGTFOIDGRSARGGCQNADCAEALNAASKADALAIIREKLPKDFIFOYHNLSNL 161

QY 61 DRI-----FDKTP 69
Db 162 DRIFAPPLEVFCPTASSFDQVPE 186
```

Search completed: January 3, 2002, 15:56:24  
Job time: 1120 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:38:48 ; Search time 65.28 seconds  
(without alignments)  
24.130 Million cell updates/sec

Title: US-09-289-346A-3  
Perfect score: 358  
Sequence: 1 TLWGEFQVGRSARGCOT.....FQPHNLNSNLDRIKFTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	218	60.9	361	US-08-838-151A-2	Sequence 2, Appli
2	218	60.9	361	US-08-838-151A-4	Sequence 4, Appli
3	218	60.9	361	US-08-838-151A-6	Sequence 6, Appli
4	218	60.9	361	US-08-838-151A-8	Sequence 8, Appli
5	217	60.6	353	US-08-838-151A-44	Sequence 44, Appli
6	217	60.6	353	US-08-838-151A-46	Sequence 46, Appli
7	217	60.6	353	US-08-838-151A-49	Sequence 49, Appli
8	217	60.6	353	US-08-838-151A-52	Sequence 52, Appli
9	217	60.6	353	US-08-838-151A-55	Sequence 55, Appli
10	212	59.2	359	US-08-809-103B-2	Sequence 2, Appli
11	212	59.2	359	US-08-809-103B-4	Sequence 4, Appli
12	212	59.2	359	US-08-809-103B-6	Sequence 6, Appli
13	212	59.2	359	US-08-809-103B-8	Sequence 8, Appli
14	190	53.1	357	US-08-838-151A-20	Sequence 20, Appli
15	190	53.1	357	US-08-838-151A-24	Sequence 24, Appli
16	190	53.1	357	US-08-838-151A-27	Sequence 27, Appli
17	190	53.1	357	US-08-838-151A-30	Sequence 30, Appli
18	63	17.6	103	US-08-552-142A-13	Sequence 13, Appli
19	60.5	16.9	131	US-08-838-151A-14	Sequence 14, Appli
20	60.5	16.9	509	US-08-737-524B-27	Sequence 27, Appli
21	59	16.5	482	US-07-792-885A-1	Sequence 1, Appli
22	59	16.5	482	US-08-142-439A-7	Sequence 7, Appli
23	59	16.5	482	US-08-869-477-7	Sequence 7, Appli
24	58.5	16.3	1713	US-08-600-982-24	Sequence 24, Appli
25	58.5	16.3	1713	PCT-US94-10261A-24	Sequence 24, Appli
26	56.5	15.8	664	US-09-295-186-17	Sequence 17, Appli
27	56	15.6	446	US-08-672-814D-11	Sequence 11, Appli

28	56	15.6	446	4	US-09-333-696-11	Sequence 11, Appli
29	55	15.4	1724	4	US-08-857-076-12	Sequence 12, Appli
30	54.5	15.2	265	4	US-09-319-989-2	Sequence 2, Appli
31	54.5	15.2	485	1	US-07-881-075-1	Sequence 1, Appli
32	54.5	15.2	485	1	US-08-120-827-1	Sequence 1, Appli
33	54.5	15.2	485	1	US-08-478-675-1	Sequence 1, Appli
34	54.5	15.2	623	4	US-09-347-801-4	Sequence 4, Appli
35	54.5	15.2	738	3	US-08-864-038A-3	Sequence 3, Appli
36	54	15.1	171	2	US-08-609-049A-22	Sequence 22, Appli
37	54	15.1	171	4	US-09-170-996-22	Sequence 22, Appli
38	54	15.1	587	2	US-08-871-266B-18	Sequence 18, Appli
39	54	15.1	587	2	US-09-018-864A-18	Sequence 18, Appli
40	54	15.1	587	3	US-08-871-267B-24	Sequence 24, Appli
41	54	15.1	587	4	US-09-618-419-24	Sequence 24, Appli
42	54	15.1	1345	2	US-08-977-767-3	Sequence 3, Appli
43	53.5	14.9	2100	2	US-08-808-793-23	Sequence 23, Appli
44	53.5	14.9	2100	3	US-08-772-512A-19	Sequence 19, Appli
45	53	14.8	258	4	US-09-251-645-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-08-838-151A-2  
; Sequence 2, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rockey, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,151A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-838-151A-2

Query Match 60.9%; Score 218; DB 4; Length 361;  
Best Local Similarity 57.1%; Pred. No. 2.4e-21;  
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TLWGEFQVGRSARGCOTSDAAAAAALQIIREKIPKYLQFPHNLNSNL 60

Db 110 TIEWGDFQIDGRSARGGQSANDSYAKALNASSVQSALAVLREEQPKDFVLQNHNRSL 169  
QY 61 DRIFDKTPEP 70  
Db 170 ERIFAKAPEP 179

RESULT 2  
US-08-838-151A-4  
; Sequence 4, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminiivirns  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,151A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; APPLICATION NUMBER: US/08/838,151A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; APPLICATION NUMBER: US/08/838,151A-4

Query Match 60.9%; Score 218; DB 4; Length 361;  
Best Local Similarity 57.1%; Pred. No. 2.4e-21;  
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDRSARGGQTSNDAAAALAAASSAAALQIIREKIPEKYLFOFHNLNSL 60  
Db 110 TIEWGDFQIDGRSARGGQSANDSYAKALNASSVQSALAVLREEQPKDFVLQNHNRSL 169  
QY 61 DRIFDKTPEP 70  
Db 170 ERIFAKAPEP 179  
RESULT 3  
US-08-838-151A-6  
; Sequence 6, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T

; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminiivirns  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,151A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; APPLICATION NUMBER: US/08-838-151A-6

Query Match 60.9%; Score 218; DB 4; Length 361;  
Best Local Similarity 57.1%; Pred. No. 2.4e-21;  
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDRSARGGQTSNDAAAALAAASSAAALQIIREKIPEKYLFOFHNLNSL 60  
Db 110 TIEWGDFQIDGRSARGGQSANDSYAKALNASSVQSALAVLREEQPKDFVLQNHNRSL 169  
QY 61 DRIFDKTPEP 70  
Db 170 ERIFAKAPEP 179

RESULT 4  
US-08-838-151A-8  
; Sequence 8, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminiivirns  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,151A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0260  
TELEPHONE: 312-616-5400  
TELEFAX: 312-616-5460  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-838-151A-8

Query Match 60.9%; Score 218; DB 4; Length 361;  
Best Local Similarity 57.1%; Pred. No. 2.4e-21;  
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDRSARGGCGTSDNDAALAAASSAAAAAQIIRKPEKYLQFQHNLSNL 60  
DB 110 TIWGFQVDRSARGGCGTSDNDAALAAASSAAAAAQIIRKPEKYLQFQHNLSNL 169  
QY 61 DRIFDKTPEP 70  
DB 170 ERIFAKAPEP 179

RESULT 5  
US-08-838-151A-44  
Sequence 44, Application US/08838151A  
Patent No. 6291743  
GENERAL INFORMATION:  
APPLICANT: Stout, John T  
APPLICANT: Luu, Hang T  
APPLICANT: Maxwell, Douglas  
APPLICANT: Ahlquist, Paul  
APPLICANT: Hanson, Steve  
TITLE OF INVENTION: Transgenic Plants Expressing Geminiiviruses  
TITLE OF INVENTION: Genes  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
STREET: Two Prudential Plaza, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,151A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0260  
TELEPHONE: 312-616-5400  
TELEFAX: 312-616-5460  
INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-838-151A-44

Query Match 60.6%; Score 217; DB 4; Length 353;  
Best Local Similarity 58.6%; Pred. No. 3.2e-21;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDRSARGGCGTSDNDAALAAASSAAAAAQIIRKPEKYLQFQHNLSNL 60  
DB 110 TIWGFQVDRSARGGCGTSDNDAALAAASSAAAAAQIIRKPEKYLQFQHNLSNL 169  
QY 61 DRIFDKTPEP 70  
DB 170 ERIFVKPEP 179

RESULT 6  
US-08-838-151A-46  
Sequence 46, Application US/08838151A  
Patent No. 6291743  
GENERAL INFORMATION:  
APPLICANT: Stout, John T  
APPLICANT: Luu, Hang T  
APPLICANT: Maxwell, Douglas  
APPLICANT: Ahlquist, Paul  
APPLICANT: Hanson, Steve  
TITLE OF INVENTION: Transgenic Plants Expressing Geminiiviruses  
TITLE OF INVENTION: Genes  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
STREET: Two Prudential Plaza, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,151A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0260  
TELEPHONE: 312-616-5400  
TELEFAX: 312-616-5460  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-838-151A-46

Query Match 60.6%; Score 217; DB 4; Length 353;  
Best Local Similarity 58.6%; Pred. No. 3.2e-21;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDRSARGGCGTSDNDAALAAASSAAAAAQIIRKPEKYLQFQHNLSNL 60  
DB 110 TIWGFQVDRSARGGCGTSDNDAALAAASSAAAAAQIIRKPEKYLQFQHNLSNL 169

QY 61 DRIFDKTPEP 70  
:|||||  
Db 170 ERIFVKVPEP 179

## RESULT 7

US-08-838-151A-49  
; Sequence 49, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
; ADDRESS: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; APPLICATION NUMBER: US/08/838,151A  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-838-151A-49

Query Match 60.6%; Score 217; DB 4; Length 353;  
Best Local Similarity 58.6%; Pred. No. 3.2e-21;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGFEQVDRSARGCGCOTSDNDAALAAASSAAALQIIREKIPEKYLFOFHNLNSNL 60  
:|||||  
Db 110 TIWGGQFQVDRSARGCGQSANDSYAKALNADSIESTILKEEQPKDYVLQHHNIRSNL 169

QY 61 DRIFDKTPEP 70  
:|||||  
Db 170 ERIFVKVPEP 179

## RESULT 8

US-08-838-151A-52  
; Sequence 52, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul

; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
; ADDRESS: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; APPLICATION NUMBER: US/08/838,151A  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 52:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-838-151A-52

Query Match 60.6%; Score 217; DB 4; Length 353;  
Best Local Similarity 58.6%; Pred. No. 3.2e-21;  
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGFEQVDRSARGCGCOTSDNDAALAAASSAAALQIIREKIPEKYLFOFHNLNSNL 60  
:|||||  
Db 110 TIWGGQFQVDRSARGCGQSANDSYAKALNADSIESTILKEEQPKDYVLQHHNIRSNL 169

QY 61 DRIFDKTPEP 70  
:|||||  
Db 170 ERIFVKVPEP 179

## RESULT 9

US-08-838-151A-55  
; Sequence 55, Application US/08838151A  
; Patent No. 6291743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
; ADDRESS: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: SVS3801P0260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-838-151A-55

```

```

Query Match 60.6%; Score 217; DB 4; Length 353;
Best Local Similarity 58.6%; Pred. No. 3.2e-21;
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAALAAASAAALQIIREKIPEKYLQFPHNLNSL 60
   :|||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 110 TIWGFQVDRSARGCGQSANDSYAKALNADSIESTLTKKEQPKDYVLOHHNIRSL 169

QY 61 DRFDKTPPEP 70
   :|||:||||
Db 170 ERIFVKVPEP 179

```

```

RESULT 10
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-809-103B-2

Query Match 59.2%; Score 212; DB 4; Length 359;
Best Local Similarity 56.5%; Pred. No. 1.5e-20;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 2 LVWGEFQVDRSARGCGQTSNDAAAALAAASAAALQIIREKIPEKYLQFPHNLNSLD 61
   :|||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 111 LEWGTQIDGRSARGCGQTANDAYAKAINAGSKSOALDVIKELAPRDYVLFHFNINSLD 170

QY 62 RIFDKTPPEP 70
   :||:||||
Db 171 KVFQVPPAP 179

```

```

RESULT 11
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-809-103B-4

```

```

Query Match 59.2%; Score 212; DB 4; Length 359;
Best Local Similarity 56.5%; Pred. No. 1.5e-20;
Matches 39; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

```











XX Hanley-Bowdoin L, Orozco BM, Kong L;  
PI WPI: 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
XX Claim 52; Page 43-44; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX Sequence 70 AA;  
SQ

Query Match 100.0%; Score 361; DB 21; Length 70;  
Best Local Similarity 100.0%; Pred. No. 7.7e-40;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60  
Db 1 TLVWGEFQVDRSGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktppep 70

RESULT 2  
AAB18677  
ID AAB18677 standard; peptide; 70 AA.  
XX  
XX AAB18677;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
DE  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Tomato golden mosaic virus.  
OS  
XX WO200054573-A1.  
PN  
XX 21-SEP-2000.  
PD  
XX 15-MAR-2000; 2000WO-US06759.  
PF  
XX 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
PA  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
PI WPI: 2000-618851/59.  
XX

PT Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
XX Disclosure; Page 18; 73pp; English.  
XX  
XX The present sequence is derived from a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
XX  
XX Sequence 70 AA;  
SQ

Query Match 95.8%; Score 346; DB 21; Length 70;  
Best Local Similarity 95.7%; Pred. No. 7.1e-38;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60  
Db 1 TLVWGEFQVDRSGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktppep 70

RESULT 3  
AAB18687  
ID AAB18687 standard; peptide; 356 AA.  
XX  
XX AAB18687;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Amino acid sequence of a geminivirus replication protein of TGMV.  
DE  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Tomato golden mosaic virus.  
OS  
XX Key Location/Qualifiers  
FH Misc-difference 354  
FT /note= "unspecified amino acid"  
ET  
XX WO200054573-A1.  
PN  
XX 21-SEP-2000.  
PD  
XX 15-MAR-2000; 2000WO-US06759.  
PF  
XX 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
PA  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
PI WPI: 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT

PT	comprise a nucleic acid construct containing a nucleic acid sequence
PT	encoding a mutant ALI protein with a mutation in the Rb binding region
PT	-
XX	
PS	Claim 53; Page 46; 73pp; English.
XX	
CC	The present sequence represents a mutant peptide, derived from a
CC	geminivirus replication (Rep) protein, also known as ALI. ALI binds
CC	double-stranded DNA, catalyzes cleavage and ligation of single-stranded
CC	DNA, and interacts with other viral and host proteins. Mutants of the
CC	protein are used to produce transgenic plants. The mutation in ALI is
CC	present in a ribosome binding region, and expression of mutant ALI
CC	protein imparts increased resistance to geminivirus infection in the
CC	plant. Mutant ALI proteins are useful for producing plants having
CC	increased resistance or reduced sensitivity to a geminivirus such as
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC	virus, cotton leaf curl virus or beet curly top virus.
XX	
SO	Sequence 70 AA:

```

Query Match          93.68; Score 338; DB 21; Length 70;
Best Local Similarity 94.33; Pred. No. 8e-37;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLVGEFQVDGRSGCGCOTSDNDAAEALNASSKEEALQIIAAAIPEKYLQFHNLSNL 60
  |||||
Db 1 tlvgefqvagrsggcqtsndaaaealnasskeaalqilirekipekylqfhnlnslnl 60
  |||||

QY 61 DRIFDKTPEP 70
  |||||
Db 61 drifdktppep 70
  |||||

RESULT 5
AAB18688
ID AAB18688 standard; peptide; 70 AA.
XX
AC AAB18688;
XX
DT 22-JAN-2001 (first entry)

```

RESULT	5	
AAAB18688		
ID	AAAB18688 standard; peptide; 70 AA.	
XX		
AC	AAAB18688;	
XX		
DT	22-JAN-2001 (first entry)	
XX		
DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.	
XX		
KW	Geminivirus; replication protein; Rep protein; AL1; transgenic plant;	
KW	ribosome binding region; resistance; geminivirus infection.	
XX		
OS	Synthetic.	
OS	Tomato golden mosaic virus.	
XX		
XX		
PH	Key	Location/Qualifiers
FT	Misc-difference 19	/note= "wild type residue replaced with Ala"
FT	Misc-difference 20	/note= "wild type residue replaced with Ala"
FT		
XX		
XX	WC200054573-Al.	
XX		
PD	21-SEP-2000.	
XX		
PF	15-MAR-2000; 2000WO-US06759.	
XX		
PR	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
XX		
XX		
PA	(UYNC-) UNIV NORTH CAROLINA STATE.	
XX		
P1	Hanley-Bowdoin L, Orozco BM, Kong L;	
XX		
DR	WPI; 2000-618851/59.	

XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant All protein with a mutation in the Rb binding region  
XX  
XX  
XX Disclosure: Page 48: 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as All. All binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the All  
CC protein are used to produce transgenic plants. The mutation in All is  
CC present in a ribosome binding region, and expression of mutant All  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant All proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX Sequence 70 AA;  
SQ  
Query Match 92.8%; Score 335; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 2e-36;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLVWGFQVDRSARGCCTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60  
|||||  
Db 1 tlwgfqvdgrsargcgcaasdaaealnasskeaalqirekipekylfqfhlnsl 60  
|||||  
QY 61 DRIFDKTPEP 70  
|||||  
Db 61 drifdktp 70  
|||||  
RESULT 6  
AAB18692  
ID AAB18692 standard; peptide: 70 AA.  
XX  
XX AAB18692;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Mutant peptide derived from amino acids 110-179 of Rep (All) protein.  
DE Geminivirus; replication protein; Rep protein; All; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
OS Tomato golden mosaic virus.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 66 /note= "wild type residue replaced with Ala"  
FT Misc-difference 69 /note= "wild type residue replaced with Ala"  
FT  
XX WO2000054573-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BW, Kong L;

XX WPI: 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant All protein with a mutation in the Rb binding region  
XX  
XX  
XX Disclosure: Page 50; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as All. All binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the All  
CC protein are used to produce transgenic plants. The mutation in All is  
CC present in a ribosome binding region, and expression of mutant All  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant All proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX Sequence 70 AA;  
SQ  
Query Match 92.5%; Score 334; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 2.7e-36;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLVWGFQVDRSARGCCTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60  
|||||  
Db 1 tlwgfqvdgrsargcgctsdnaaealnasskeaalqirekipekylfqfhlnsl 60  
|||||  
QY 61 DRIFDKTPEP 70  
|||||  
Db 61 drifdktp 70  
|||||  
RESULT 7  
AAB18684  
ID AAB18684 standard; peptide: 70 AA.  
XX  
XX AAB18684;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Mutant peptide derived from amino acids 110-179 of Rep (All) protein.  
DE Geminivirus; replication protein; Rep protein; All; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.  
OS Tomato golden mosaic virus.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 7 /note= "wild type residue replaced with Ala"  
FT Misc-difference 8 /note= "wild type residue replaced with Ala"  
FT  
XX WO2000054573-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX PA

(UYNC-) UNIV NORTH CAROLINA STATE.  
Hanley-Bowdoin L, Orozco BM, Kong L;  
WPI: 2000-618951/59.  
Transgenic plants with increased resistance to geminivirus infection  
comprise a nucleic acid construct containing a nucleic acid sequence  
encoding a mutant All protein with a mutation in the Rb binding region  
disclosure; Page 49; 73pp; English.  
The present sequence represents a mutant peptide, derived from a  
geminivirus replication (Rep) protein, also known as All. All binds  
double-stranded DNA, catalyses cleavage and ligation of single-stranded  
DNA, and interacts with other viral and host proteins. Mutants of the  
protein are used to produce transgenic plants. The mutation in All is  
present in a ribosome binding region, and expression of mutant All  
protein imparts increased resistance to geminivirus infection in the  
plant. Mutant All proteins are useful for producing plants having  
increased resistance or reduced sensitivity to a geminivirus such as  
tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf  
curl virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
virus, cotton leaf curl virus or beet curly top virus.

[illegible]

DB		of drifakcpep 70
	RESULT	9
	AAB18678	
ID	AAB18678	standard; peptide; 70 AA.
XX	AC	
XX	AAB18678;	
XX	22-JAN-2001	(first entry)
XX	Mutant peptide derived from amino acids 110-179 of Rep (All) protein.	
DE	Geminivirus; replication protein; Rep protein; All; transgenic plant;	
XX	ribosome binding region; resistance; geminivirus infection.	
KW	Synthetic.	
OS	Tomato golden mosaic virus.	
XX	Key	Location/Qualifiers
PH	Misc-difference 12	/note= "wild type residue replaced with Ala"
FT	Misc-difference 13	/note= "wild type residue replaced with Ala"
FT	Misc-difference 15	/note= "wild type residue replaced with Ala"
FT	Misc-difference 15	/note= "wild type residue replaced with Ala"
FT	Misc-difference 15	/note= "wild type residue replaced with Ala"
XX	WO200054573-Al.	
PV	21-SEP-2000.	
XX		
XX		
XX		







```

OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 47 /note= "wild type residue replaced with Ala"
FT Misc-difference 48 /note= "wild type residue replaced with Ala"
FT Misc-difference 49 /note= "wild type residue replaced with Ala"
FT
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
PS Claim 52; Page 44; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;
Query Match 90.0%; Score 325; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 4e-35;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLFFQFHNLSNL 60
Db 1 tlvwgefvdgrsgrgcgtsndaaaealnasskeaalqirekipekylffqfhnlnsal 60
QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70
RESULT 14
AAB18683
ID AAB18683 standard; peptide: 70 AA.
XX
AC AAB18683;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX

```

---

```

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 59 /note= "wild type residue replaced with Ala"
FT Misc-difference 61 /note= "wild type residue replaced with Ala"
FT Misc-difference 62 /note= "wild type residue replaced with Ala"
FT
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI: 2000-618851/59.
XX
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX
PS Claim 53; Page 45; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;
Query Match 89.8%; Score 324; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 5.4e-35;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TLVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQITAAAIPEKYLFFQFHNLSNL 60
Db 1 tlvwgefvdgrsgrgcgtsndaaaealnasskeaalqirekipekylffqfhnlnsal 60
QY 61 DRIFDKTPEP 70
Db 61 aaifdktp 70
RESULT 15
AAB18682
ID AAB18682 standard; peptide: 70 AA.
XX
AC AAB18682;
XX

```

Job time: 154 sec

DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 52 /note= "wild type residue replaced with Ala"  
FT Misc-difference 54 /note= "wild type residue replaced with Ala"  
FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
FT  
XX  
PN WO200054573-A1.  
XX  
XX  
PD 21-SEP-2000.  
XX  
XX  
PF 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
PA  
XX  
PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX WPI; 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
XX Claim 53; Page 44-45; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
XX Sequence 70 AA;

Query Match 89.2%; Score 322; DB 21; Length 70;  
Best Local Similarity 91.4%; Pred. No. 9.9e-35;  
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 TLVWGFQVQDGRSARGGCGTSDNDAEALNASSKEEALQITIAAAIPEKYLFPFHNLNSL 60  
Db 1 tlvwgfqvqdgrrsarggqtsndaaalnasskecalqiirekipekylfafaalnsnl 60  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktppep 70



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:40:08 ; Search time 72.79 Seconds  
(without alignments)  
73.255 Million cell updates/sec

Title: US-09-289-346A-4

Perfect score: 361

Sequence: 1 TLVWGEFQVDRSARGCQT.....FQFNLNSLDRIFDKTPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	346	95.8	352	1 QOCVLI	ALI protein - toma
2	249	69.0	361	1 QOCVPT	ALI protein - pota
3	240	66.5	358	2 S07594	hypothetical prote
4	233	64.5	362	1 J01887	ALI protein - toma
5	229	63.4	359	2 S39211	gene C1 protein -
6	222	61.5	349	2 J02300	replicase - pepper
7	222	61.5	349	2 S31875	ALI protein - pepp
8	222	61.5	359	2 S22593	hypothetical prote
9	219	60.7	359	2 S39235	gene C1 protein -
10	216	59.8	351	2 J02327	ALI protein - indi
11	216	59.8	355	1 QOCVW1	AV1 protein - abut
12	213	59.0	385	2 S28360	ALI protein - beet
13	210	58.2	358	1 J01870	ALI protein - toma
14	200	55.4	357	1 QOCVCI	ALI protein - toma
15	199	55.1	360	2 S39885	replication-associ
16	134	37.1	131	2 S45059	AC1 protein (clone
17	128	35.5	347	1 QOCVS1	ALI protein - squa
18	72	19.9	587	2 JG1419	Fc gamma (IgG) rec
19	63.5	17.6	584	2 T19061	hypothetical prote
20	62.5	17.3	1006	2 T34992	hypothetical prote
21	61.5	17.0	299	2 B71967	probable peptidyl-
22	61.5	17.0	335	1 DBBSGF	glyceroldehyde-3-p
23	61.5	17.0	481	2 A70091	probable phosphoe
24	61.5	17.0	1044	2 T43800	protein kinase bub
25	61.5	17.0	1203	2 T17415	mycelial surface a
26	61.5	17.0	1229	2 T48959	kinesin-like prote
27	60.5	16.8	159	2 C71838	NADH dehydrogenase
28	60.5	16.8	159	2 E64677	NADH dehydrogenase
29	60	16.6	767	2 F71479	hypothetical prote

30	60	16.6	1008	2 T41244	SEC14 protein homo
31	59	16.3	160	2 G82060	hypothetical prote
32	59	16.3	1502	1 RGBYH1	CYC1/CYP3 transcri
33	58.5	16.2	297	1 MNVNRV	nonstructural prot
34	58.5	16.2	357	2 T02246	hypothetical prote
35	58.5	16.2	769	2 F81742	conserved hypotet
36	58.5	16.2	2135	2 T14602	variant-specific s
37	58	16.1	555	2 C45868	glycerol-3-phospha
38	58	16.1	601	2 F84979	sulfite reductase
39	57.5	15.9	858	2 C64577	virB4 homolog He
40	57.5	15.9	1409	2 S74916	alkaline phosphata
41	57	15.8	152	2 C96602	hypothetical prote
42	57	15.8	224	2 B56168	deoxyguanosine kin
43	57	15.8	380	2 S70594	ubiquinol--cytochr
44	57	15.8	604	2 T19682	hypothetical prote
45	57	15.8	771	2 B38252	granulocyte colony

## ALIGNMENTS

### RESULT 1

QOCVLI

ALI protein - tomato golden mosaic virus

C:Species: tomato golden mosaic virus

A:Note: host Nicotiana sp. (tobacco)

C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994

C:Accession: A04170

R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.

EMBO J. 3, 2197-2205, 1984

A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma

A:Reference number: A04163

A:Accession: A04170

A:Molecule type: DNA

A:Residues: 1-352 <HAM>

C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.8%; Score 346; DB 1; Length 352;

Best Local Similarity 95.7%; Pred. No. 3e-33;

Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLSNL 60

|||||

DB 111 TLVWGEFQVDRSARGCQTSNDAAAEALNASKEEALQIIREKIPEKYLFOFHNLSNL 170

QY 61 DRIFDKTPEP 70

|||||

DB 171 DRIFDKTPEP 180

### RESULT 2

QOCVPT

ALI protein - potato yellow mosaic virus (isolate Venezuela)

C:Species: potato yellow mosaic virus

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C:Accession: JU0364

R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.

J. Gen. Virol. 72, 1515-1520, 1991

A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye

A:Reference number: JU0362; MUID:91311403

A:Accession: JU0364

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-361 <COU>

A:Cross-references: GB:D00940; NID:q222458; PIDN:BAA00782.1; PID:q222459

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

```
Query Match          69.0%; Score 249; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 9.4e-22;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCGTSDNDAAEALNASSKEALQIIAAAIPEKYLFOFHNLNSNL 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TLVWGLFOIDGRSARGGQTVNDAAAEALNSGTKEANKIIEKLPKFLFOYHNLSCNL 169

QY 61 DRIFDKTPE 69
   ||||| ||
Db 170 DRIFMKAPE 178

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C:Species: cassava latent virus
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: S07594
R:Morris, B.; Coates, L.; Love, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07590; MUID:90174930
A:Accession: S07594
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <MOR>
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C:Genetics:
A:Map position: segment DNAL
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match          66.5%; Score 240; DB 2; Length 358;
Best Local Similarity 61.4%; Pred. No. 1.1e-20;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCGTSDNDAAEALNASSKEALQIIAAAIPEKYLFOFHNLNSNL 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 109 TVEWGQFQIDGRSARGGQSDANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLSNL 168

QY 61 DRIFDKTPEP 70
   ||||| :||
Db 169 DRIFQEPAP 178

RESULT 4
JQ1887
A1 protein - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: C1 protein
C:Species: tomato yellow leaf curl virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: JQ1887
R:Dray, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: JQ1885; MUID:93139778
A:Accession: JQ1887
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DRY>
A:Cross-references: GB:S53251
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match          64.5%; Score 233; DB 1; Length 362;
Best Local Similarity 54.1%; Pred. No. 7.4e-20;
Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDRSGRGCGTSDNDAAEALNASSKEALQIIAAAIPEKYLFOFHNLNSNL 60
   || |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
Db 110 TLVWGEFQIDGRSARGGQSDANDAYAQALNTGSKSEALNVRELAPKDYVLFQFHNLSNL 169

QY 61 DRI-----FDKTPPE 69
   ||| |||
Db 170 DRIFTPLEVYVSPFLSSSFDKRVPE 194

RESULT 5
S39211
gene C1 protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C:Accession: S39211
R:Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A:Description: High similarity among the tomato yellow leaf curl virus isolates from
A:Reference number: S39209
A:Accession: S39211
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <NOR>
A:Cross-references: EMBL:225751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match          63.4%; Score 229; DB 2; Length 359;
Best Local Similarity 60.9%; Pred. No. 2.2e-19;
Matches 42; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDRSGRGCGTSDNDAAEALNASSKEALQIIAAAIPEKYLFOFHNLNSNL 61
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 LEWCTFQIDGRSARGGQTDANDAYAKAINAGSKSEALDYIKELAPDYILHFNHNSNL 170

QY 62 DRIFDKTPEP 70
   ||| |||
Db 171 RVFQVPPAP 179

RESULT 6
JQ2300
replicase - pepper huasteco virus (component A)
N:Alternate names: ORF ALL protein
C:Species: pepper huasteco virus
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: JQ2300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: JQ2299; MUID:94015007
A:Accession: JQ2300
A:Molecule type: DNA
A:Residues: 1-349 <TOR>
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match          61.5%; Score 222; DB 2; Length 349;
Best Local Similarity 60.0%; Pred. No. 1.4e-18;
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGRGCGTSDNDAAEALNASSKEALQIIAAAIPEKYLFOFHNLNSNL 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 TVEWGEFQIDGRSARGGQSDANDTYAKALNSASAEALQIIKEEQPOHFFLQFHNIVSNA 169

QY 61 DRIFDKTPEP 70
   |||| |||
Db 170 NRIFQTPPEP 179

RESULT 7
S31875
A1 protein - pepper rizado amarillo virus
C:Species: pepper rizado amarillo virus
```

C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999

C:Accession: S31875  
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu

submitted to the EMBL Data Library, February 1993

A:Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa

A:Reference number: S31872

A:Accession: S31875

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-349 <TOR>

A:Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025

A:Note: the source is designated as pepper huasteco virus

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.5%; Score 222; DB 2; Length 349;  
Best Local Similarity 60.0%; Pred. No. 1.4e-18;  
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLNSNL 60

Db 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLNSNL 60

Db 110 TVWGEFQIDGRSARGGQSQANDYAKALNSASAEALQIIKEEPOHFFLQFHNLVNSNA 169

QY 61 DRIFDKTPEP 70

Db 170 NRIFQTPPEP 179

RESULT 8

S22593

hypothetical protein C4 - tomato yellow leaf curl virus

C:Species: tomato yellow leaf curl virus

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999

C:Accession: S22593

R:Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.

Nucleic Acids Res. 19, 6763-6769, 1991

A>Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartit

A:Reference number: S22588; MUID:92107660

A:Accession: S22593

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-359 <KHE>

A:Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.5%; Score 222; DB 2; Length 359;  
Best Local Similarity 58.0%; Pred. No. 1.5e-18;  
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDRSARGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLNSNL 61

Db 111 LEWGTQFQIDGRSARGGQQTANDAYAKAINAGSKSQAOLDVIKELAPRDYVLFHFHNTSSNLD 170

QY 62 RIFDKTPEP 70

Db 171 KVFQPPAP 179

RESULT 9

S39235

gene C1 protein - tomato yellow leaf curl virus

C:Species: tomato yellow leaf curl virus

C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Sep-1997

C:Accession: S39235

R:Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.

submitted to the EMBL Data Library, December 1993

A:Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.

A:Reference number: S39233

A:Accession: S39235

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-359 <CRE>

A:Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.7%; Score 219; DB 2; Length 359;  
Best Local Similarity 58.0%; Pred. No. 3.3e-18;  
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDRSARGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLNSNL 61

Db 111 LEWGTQFQIDGRSARGGQQTANDAYAKAINARSKSEALDIKOLAPRDYVLFHFHNTSSNLD 170

QY 62 RIFDKTPEP 70

Db 171 KVFQPPAP 179

RESULT 10

JQ2327

AL1 protein - Indian cassava mosaic virus

N:Alternate names: replication-associated protein

C:Species: Indian cassava mosaic virus

C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999

C:Accession: JQ2327; S35883

R:Hong, Y.G.; Robinson, D.J.; Harrison, B.D.

J. Gen. Virol. 74, 2437-2443, 1993

A>Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t

A:Reference number: JQ2326; MUID:94065670

A:Accession: JQ2327

A:Molecule type: DNA

A:Residues: 1-351 <RON>

A:Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 59.8%; Score 216; DB 2; Length 351;  
Best Local Similarity 61.2%; Pred. No. 7.3e-18;  
Matches 41; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 4 WGEFQVDRSARGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLNSNL 63

Db 113 WGTQFQIDGRSARGGQSQANDAYAAALNSGSKSEALKILRELAPRDYVLFHFHNTSSNLD 172

QY 64 FDKTPEP 70

Db 173 FTKPPPP 179

RESULT 11

QOCVW1

AVI protein - abutilon mosaic virus (isolate West India)

C:Species: abutilon mosaic virus

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994

C:Accession: A36214

R:Frischmuth, T.; Zimmatt, G.; Jeske, H.

Virolgy 178, 461-468, 1990

A>Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as

A:Reference number: A36214; MUID:91020984

A:Accession: A36214

A:Molecule type: DNA

A:Residues: 1-355 <PRI>

A:Cross-references: EMBL:X15983

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 59.8%; Score 216; DB 1; Length 355;  
Best Local Similarity 58.6%; Pred. No. 7.4e-18;  
Matches 41; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLNSNL 60

```

Db 110 TAWEQFQIDGRSARGGQQTANDSYAKALNAGDVQSALNLTKEEQPKDYVLQNHNIRSNL 169
      :||||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 DRIFDKTPEP 70
      :|||:|
Db 170 ERIFAKAPEP 179
      :|||:|

RESULT 12
S28360
ALI protein - beet curly top virus
C:Species: beet curly top virus
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C:Accession: S28360
R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly top
A:Reference number: S28360
A:Accession: S28360
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-385 <STA>
A:Cross-references: GB:IM24597; EMBL:X04144; NID:9210678; PIDN:AAA42751.1; PID:g210679
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.0%; Score 213; DB 2; Length 385;
Best Local Similarity 55.7%; Pred. No. 1.8e-17;
Matches 39; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGGCGTSDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLSNL 60
      :||||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 137 TIWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196
      :|||:|

QY 61 DRIFDKTPEP 70
      :||:|
Db 197 QKIFRPPDP 206
      :||:|

RESULT 13
JQ1870
ALI protein - tomato mottle virus (isolate Florida)
C:Species: tomato mottle virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: JQ1870
R:Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A:Reference number: JQ1870; MUID:93107858
A:Accession: JQ1870
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <ABO>
A:Cross-references: GB:L14460
C:Genetics:
A:Map position: segment A
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 58.2%; Score 210; DB 1; Length 358;
Best Local Similarity 54.3%; Pred. No. 3.9e-17;
Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDGRSARGGCGTSDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLSNL 60
      :||||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 107 TIWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALAVLREEQPKDYVLQNHNIRSNL 166
      :|||:|

QY 61 DRIFDKTPEP 70
      :|||:|
Db 167 ERIFAKAPEP 176
      :|||:|

RESULT 14
```

```

QCCVC1
ALI protein - tomato yellow leaf curl virus
N:Alternate names: Cl protein
C:Species: tomato yellow leaf curl virus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: D40779
R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virolgy 185, 151-161, 1991
A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A:Reference number: A40779; MUID:92024070
A:Accession: D40779
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-357 <NAV>
A:Cross-references: GB:X15656; NID:962204; PIDN:CAA33688.1; PID:g62207
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 55.4%; Score 200; DB 1; Length 357;
Best Local Similarity 63.9%; Pred. No. 5.9e-16;
Matches 39; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGEFQVDGRSARGGCGTSDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLSNLDRI 63
      :||||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 111 FGVSQIDGRSARGGQGSANDAYAEALNSGSKSEALNLTKEAPKDYILQFHNLSNLDRI 170
      :||:|

QY 64 F 64
Db 171 F 171

RESULT 15
S59885
replication-associated protein Cl - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C>Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C:Accession: S59885
R:Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A:Reference number: S58346
A:Accession: S59885
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <HON>
A:Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA89229.1; PID:g974211
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 55.1%; Score 199; DB 2; Length 360;
Best Local Similarity 57.6%; Pred. No. 7.8e-16;
Matches 38; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 4 WGEFQVDGRSARGGCGTSDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLSNLDRI 63
      :||||:|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 113 FGVFQIDGRSARGGQGSANDAYAEALNSGSKAAALDILREKAPKDYFLQFHNLSNLDRI 172
      :||:|

QY 64 FDKTPE 69
      :|:|
Db 173 FTPSAE 178
      :|:|

Search completed: January 3, 2002, 15:40:09
Job time: 305 sec
```





---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 3, 2002, 15:57:15 ; Search time 43.68 Seconds  
(without alignments)  
58.758 Million cell updates/sec

Title: US-09-289-346A-4  
Perfect score: 361  
Sequence: 1 TLVWGEFQVDRSARGCQT.....FQFHNLNSLDRIFDKTPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	346	95.8	352	1 VAL1_TGMV	P03567 tomato gold
2	249	69.0	361	1 VAL1_PYMV	P27258 potato yell
3	240	66.5	358	1 VAL1_CLVK	P14982 cassava lat
4	240	66.5	358	1 VAL1_CLVN	P14972 cassava lat
5	233	64.5	362	1 VAL1_TYLC	P36279 tomato yell
6	229	63.4	359	1 VAL1_TYLCU	P38609 tomato yell
7	222	61.5	349	1 VAL1_PHUV	Q08923 pepper huas
8	222	61.5	359	1 VAL1_TYLCM	P27260 tomato yell
9	217	60.1	353	1 VAL1_BGMV	P05175 bean golden
10	216	59.8	355	1 VAL1_ABMVW	P21947 abutilon mo
11	213	59.0	358	1 VAL1_BCTV	P14991 beet curly
12	210	58.2	361	1 VAL1_TMOV	Q06657 tomato moft
13	200	55.4	357	1 VAL1_TYLCV	P27259 tomato yell
14	128	35.5	347	1 VAL1_SLICV	P29048 squash leaf
15	61.5	17.0	299	1 Y175_HELPJ	Q92mq7 helicobacte
16	61.5	17.0	334	1 G3P_BACST	P00362 bacillus st
17	61.5	17.0	335	1 G3P_BACCO	P15115 bacillus co
18	61.5	17.0	1044	1 BUB1_SCHPO	Q94751 schizosacch
19	59	16.3	1483	1 CYP1_YEAST	P12351 saccharomyc
20	58.5	16.2	297	1 RRPB_RABVP	Q06747 rabies viru
21	58	16.1	492	1 MOT3_MOUSE	Q35308 mus musculu
22	58	16.1	492	1 MOT3_RAT	Q70461 rattus norv
23	58	16.1	555	1 GLPD_BACSU	P18158 bacillus su
24	58	16.1	601	1 CYS1_BUCAL	P57503 buchnera ap
25	57	15.8	223	1 DGR2_LACAC	Q59484 lactobacill
26	57	15.8	316	1 SPDE_COFAR	Q82147 coifera arab
27	57	15.8	379	1 CYP1_ASTPE	Q33818 asterina pe
28	57	15.8	387	1 Y4PF_RHISN	P55615 rhizobium s
29	57	15.8	836	1 GCSR_HUMAN	Q99062 homo sapien
30	57	15.8	1287	1 SKI2_YEAST	P35207 saccharomyc
31	56.5	15.7	863	1 AMFN_CAUCR	P37893 caulobacter
32	56	15.5	367	1 LHX4_MOUSE	P53776 mus musculu
33	56	15.5	577	1 PTLB_LACCA	P24400 lactobacill

RESULT 1

ID	VAL1_TGMV	STANDARD;	PRT;	352 AA.
AC	P03567;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	01-JUN-1994 (Rel. 29, Last annotation update)			
DE	AL1 PROTEIN.			
GN	AC1.			
OS	Tomato golden mosaic virus (TGMV).			
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.			
OX	NCHI_taxid=10831;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;			
RT	*Complete nucleotide sequence of the infectious cloned DNA components			
RT	of tomato golden mosaic virus: potential coding regions and regulatory			
RT	sequences.;			
RL	EMBO J. 3:2197-2205(1984).			
CC	-!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; K02029; -; NOT_ANNOTATED_CDS.			
DR	PIR; A04170; Q0CVLL.			
DR	InterPro; IPR001191; Gemini_AL1.			
DR	Pfam; PF00799; Gemini_AL1; 1.			
DR	PRINTS; PR00227; GEMCOATALL.			
DR	PRINTS; PR00228; GEMCOATCLVL.			
DR	ProDom; PD000736; Gemini_AL1; 1.			
KW	ATP-binding. 223 230 ATP (POTENTIAL).			
FT	NP_BIND 352 AA; 40332 MW; C33C938E964B4A4 CRG64;			
SQ	SEQUENCE			

Query Match 95.8%; Score 346; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 1.le-33;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	TLVWGEFQVDRSARGCQT	SDNDAAEALNASSKEALQIIAAAIPEKYLQFQHNLSNL	60
DB	111	TLVWGEFQVDRSARGCQT	SDNDAAEALNASSKEALQIIREKIPEKYLQFQHNLSNL	170

QY	61	DRIFDKTPEP	70
DB	171	DRIFDKTPEP	180

RESULT 2

VAL1\_PYMV

ID VAL1\_PYMVV STANDARD; PRT; 361 AA.  
AC P27258;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE AL1 PROTEIN.  
OS Potato yellow mosaic virus (isolate Venezuela).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10828;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91311403; PubMed=1856690;  
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;  
RT "The nucleotide sequence of the infectious cloned DNA components of  
RT potato yellow mosaic virus."  
RL J. Gen. Virol. 72:1515-1520(1991).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; D00940; BAA00782.1; --  
DR PIR; J00364; QOCVPT.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLI.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
KW ATP-binding. 222 229 ATP (POTENTIAL).  
FT NP\_BIND  
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match 69.0%; Score 249; DB 1; Length 361;  
Best Local Similarity 68.1%; Pred. No. 3.6e-22;  
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDSARGCGQTSNDAAAEALNASSKEALQIIAAAIPEKYLQFPHNLNSNL 60  
DB 110 TIWGLFQIDGRSARGCGQTVNDAAAEALNASSKEALQIIAAAIPEKYLQFPHNLNSNL 169  
QY 61 DRIFDKTPE 69  
DB 170 DRIFDKAPE 178

RESULT 3  
VAL1\_CLVK STANDARD; PRT; 358 AA.  
AC P14982;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 PROTEIN (40.4 KDA PROTEIN).  
GN AC1.  
OS Cassava latent virus (strain West Kenya 844).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stanley J., Gay M.R.;  
RT "Nucleotide sequence of cassava latent virus DNA."  
RL Nature 301:260-262(1983).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; J02057; --; NOT\_ANNOTATED\_CDS.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLI.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
KW ATP-binding. 220 227 ATP (POTENTIAL).  
FT NP\_BIND  
SQ SEQUENCE 358 AA; 40346 MW; ED173E753E92D69 CRC64;

Query Match 66.5%; Score 240; DB 1; Length 358;  
Best Local Similarity 61.4%; Pred. No. 4.2e-21;  
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDSARGCGQTSNDAAAEALNASSKEALQIIAAAIPEKYLQFPHNLNSNL 60  
DB 109 TVEWGFQIDGRSARGCGQSANDAYAKALNSGSKSEALNVIRELVKDFVLQPHNLNSNL 168  
QY 61 DRIFDKTPEP 70  
DB 169 DRIFQEPAP 178

RESULT 4  
VAL1\_CLVN STANDARD; PRT; 358 AA.  
AC P14972;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 PROTEIN (40.4 KDA PROTEIN).  
GN AC1.  
OS Cassava latent virus (strain Nigerian).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10819;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90174930; PubMed=2308631;  
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;  
RT "Nucleotide sequence of the infectious cloned DNA components of  
RT African cassava mosaic virus (Nigerian strain)."  
RL Nucleic Acids Res. 18:197-198(1990).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; X17095; CAA34953.1; --  
DR PIR; S07594; S07594.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLI.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
KW ATP-binding. 220 227 ATP (POTENTIAL).  
FT NP\_BIND  
SQ SEQUENCE 358 AA; 40435 MW; 1DB16BB0CB2D5E2C CRC64;

Query Match 66.5%; Score 240; DB 1; Length 358;  
Best Local Similarity 61.4%; Pred. No. 4.2e-21;  
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDSARGCGQTSNDAAAEALNASSKEALQIIAAAIPEKYLQFPHNLNSNL 60  
DB 109 TVEWGFQIDGRSARGCGQSANDAYAKALNSGSKSEALNVIRELVKDFVLQPHNLNSNL 168  
QY 61 DRIFDKTPEP 70  
DB 169 DRIFQEPAP 178

```

RT *Complete nucleotide sequence of pepper huasteco virus: analysis and
RT comparison with bipartite geminiviruses.";
RL J. Gen. Virol. 74:3225-2231(1993).
CC -----
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL: X70418; CAA49856.1; -.
DR PIR: S31875; S31875.
DR PIR: JQ2300; JQ2300.
DR InterPro: IPR001191; Gemini_ALL.
DR Pfam: PF00799; Gemini_ALL; 1.
DR PRINTS: PR00227; GEMCOATALL1.
DR PRINTS: PR00228; GEMCOATCLV1.
DR -----

```



```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15983; -: NOT_ANNOTATED_CDS.
DR PIR; A36214; OOCVW1.
DR InterPro; IPR001191; Gemini_ALL.
DR Pfam; PF00799; Gemini_ALL; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLI.
DR ProDom; PD000736; Gemini_ALL; 1.
KW ATP-binding.
FT NP_BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 355 AA; 40257 MW; 16A2CABA63251E95 CRC64;

Query Match 59.8%; Score 216; DB 1; Length 355;
Best Local Similarity 58.6%; Pred. No. 3e-18; Mismatches 19; Indels 0; Gaps 0;
Matches 41; Conservative 10;

QY 1 TLVWGEFQDGRSARGGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60
DB 110 TAEWGEFQIDGRSARGGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 169

QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179

RESULT 11
VAL1_BCTV
ID VAL1_BCTV STANDARD; PRT; 358 AA.
AC P14991;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE ALL PROTEIN (40.8 KDA PROTEIN).
OS Beet curly top virus (BCTV).
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX NCBI_TaxID=10840;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT "The nucleotide sequence of an infectious clone of the geminivirus
RT beet curly top virus."
RL EMBO J. 5:1761-1767(1986).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04144; -: NOT_ANNOTATED_CDS.
DR InterPro; IPR001191; Gemini_ALL.
DR Pfam; PF00799; Gemini_ALL; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLI.
DR ProDom; PD000736; Gemini_ALL; 1.
KW ATP-binding.
FT NP_BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 59.0%; Score 213; DB 1; Length 358;
```

```
Best Local Similarity 55.7%; Pred. No. 6.8e-18;
Matches 39; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQDGRSARGGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60
DB 110 TAEWGEFQIDGRSARGGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 169

QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179

RESULT 12
VAL1_TMOV
ID VAL1_TMOV STANDARD; PRT; 361 AA.
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN ALL.
OS Tomato mottle virus (isolate Florida) (TMov).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=36449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107858; PubMed=1469361;
RA Abouzid A.M., Polston J.E., Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT isolated from tomatoes in Florida."
RL J. Gen. Virol. 73:3225-3229(1992).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14460; AAC32414.1; -.
DR PIR; JQ1870; JQ1870.
DR InterPro; IPR001191; Gemini_ALL.
DR Pfam; PF00799; Gemini_ALL; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLI.
DR ProDom; PD000736; Gemini_ALL; 1.
KW ATP-binding.
FT NP_BIND 222 229 ATP (BY SIMILARITY).
SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEEAC6950 CRC64;

Query Match 58.2%; Score 210; DB 1; Length 361;
Best Local Similarity 54.3%; Pred. No. 1.6e-17;
Matches 38; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQDGRSARGGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60
DB 110 TAEWGEFQIDGRSARGGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 169

QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179

RESULT 13
VAL1_TYLCV
ID VAL1_TYLCV STANDARD; PRT; 357 AA.
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
```

```
DE AL1 PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024070; PubMed=1926771;
RA Navot N., Pichersky E., Zeldan M., Zamir D., Czosnek H.;
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
with a single genomic component.";
RL Virology 185:151-161(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X15656; CAA33688.1; -.
DR PIR; D40779; QOCVCL.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLL.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 219 226 ATP (POTENTIAL).
SQ SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;

Query Match 55.4%; Score 200; DB 1; Length 357;
Best Local Similarity 63.9%; Pred. No. 2.4e-16;
Matches 39; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGEFQVDRSGARGCQTSNDAAAEALNASSKEKALQITAAAIPEKYLFOFHNLSNLDRI 63
DB 111 FGVSGIDGRSGARGQOOSANDAYAEALNAGSKSEALNKEKAPKDYILOFHNLSNLDRI 170

QY 64 F 64
DB 171 F 171

RESULT 14
VAL1_SLCV
ID VAL1_SLCV STANDARD; PRT; 347 AA.
AC P29048.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE AL1 PROTEIN.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082449; PubMed=1984668;
RA Lazarowitz S.G., Lazdins I.B.;
RT "Infectivity and complete nucleotide sequence of the cloned genomic
components of a bipartite squash leaf curl geminivirus with a broad
host range phenotype.";
RL Virology 180:58-69(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M38183; AAC32410.1; ALT_INIT.
DR PIR; C36785; QOCVSL.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLL.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 218 225 ATP (POTENTIAL).
SQ SEQUENCE 347 AA; 39110 MW; AFDAEBDD122110E CRC64;

Query Match 35.5%; Score 128; DB 1; Length 347;
Best Local Similarity 39.4%; Pred. No. 8.2e-08;
Matches 26; Conservative 13; Mismatches 23; Indels 4; Gaps 1;

QY 5 GEFQVDRSGARGCQTSNDAAAEALNASSKEKALQITAAAIPEKYLFOFHNLSNLDRI 64
DB 116 GQYKVGSG-----GSKSNKDDVYHNAVAGSAGEALDIKAGDPRTFIVYHNLNLANVERUF 171

QY 65 DKTPEP 70
DB 172 QKPPEP 177

RESULT 15
Y175_HELPJ
ID Y175_HELPJ STANDARD; PRT; 299 AA.
AC Q92M07.
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN JHP0161 PRECURSOR.
GN JHP0161.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC STRONG, TO C.JEJUNI CBF2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001454; AAD05744.1; -.
DR HSSP; Q13526; IPIN.
DR InterPro; IPR000297; Rotamase.
DR Pfam; PF00639; Rotamase; 1.
DR PROSITE; PS01096; PPIC_PPIASE_1; 1.
KW Hypothetical protein; Isomerase; Rotamase; Signal; Complete proteome.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 299 HYPOTHETICAL PROTEIN JHP0161.
FT DOMAIN 154 253 PPIC-LIKE.
SQ SEQUENCE 299 AA; 34040 MW; 9C037B1CD110143 CRC64;
```





---

**THIS PAGE BLANK (USPTO)**



```
O09727 ID O09727 PRELIMINARY; PRT; 226 AA.
AC Q09727;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR InterPro; IPR001191; Gemin_A1.
DR Pfam; PF00799; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_A1; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 77.8%; Score 281; DB 12; Length 226;
Best Local Similarity 78.6%; Pred. No. 3.1e-25;
Matches 55; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGFQVDSRGSGCGTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60
Db 111 TLVWGFQVDSRGSGCGTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 170
QY 61 DRIFDKTPEP 70
Db 171 DRIFAKAPEP 180

RESULT 3
Q9WHF6 ID Q9WHF6 PRELIMINARY; PRT; 226 AA.
AC Q9WHF6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN96-H5KW;
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M., Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131071; AAD33471.1; -.
DR InterPro; IPR001191; Gemin_A1.
DR Pfam; PF00799; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_A1; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 77.8%; Score 281; DB 12; Length 226;
Best Local Similarity 74.3%; Pred. No. 3.1e-25;
Matches 52; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVWGFQVDSRGSGCGTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60
Db 111 TLVWGFQVDSRGSGCGTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 170
```

```
Db 111 TLVWGFQVDSRGSGCGTSDNDAAEALNASSKEEAMRIIEKLPKEFLFOYHNLSSNL 170
QY 61 DRIFDKTPEP 70
Db 171 DRIFAKAPEP 180

RESULT 4
Q67574 ID Q67574 PRELIMINARY; PRT; 361 AA.
AC Q67574;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE REPLICATIVE PROTEIN.
GN A1.
OS Bean golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10839;
RN [1]
RP SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G., Maxwell D.P., Russell D.R.;
RL Phytopathology 81:980-985(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C., Morales F.J., Maxwell D.P.;
RL Plant Dis. 75:336-342(1991).
RN [3]
RP SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M88686; AAA46312.1; -.
DR InterPro; IPR001191; Gemin_A1.
DR Pfam; PF00799; Gemin_A1; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_A1; 1.
SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.3%; Score 279; DB 12; Length 361;
Best Local Similarity 77.6%; Pred. No. 9e-25;
Matches 52; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 4 WGEFQVDSRGSGCGTSDNDAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSLDRI 63
Db 113 WGEFQVDSRGSGCGTSDNDAAEALNASSKEEAMQIIEKLPKEFLFOYHNLSSNL 172
QY 64 DRIFDKTPEP 70
Db 173 DRIFAKAPEP 179

RESULT 5
Q9QDB1 ID Q9QDB1 PRELIMINARY; PRT; 225 AA.
AC Q9QDB1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS cowpea golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=69263;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGMV-BR;
RA Faria J.C.;
RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from Brazil.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
```

DR EMBL; AF188708; AAF06318.1; .  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
FT NON\_TER 225 225  
SQ SEQUENCE 225 AA; 25766 MW; 1089C86BD8D15B5D CRC64;

Query Match 76.7%; Score 277; DB 12; Length 225;  
Best Local Similarity 77.6%; Pred. No. 9e-25; Mismatches 8; Indels 0; Gaps 0;  
Matches 52; Conservative 7;

QY 4 MGFEFQVDRSARGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFPQHNLNSNLDRI 63  
DB 113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPKFLFYHNLSSNLDRI 172

QY 64 FDKTPEP 70

DB 173 FKPPPEP 179

RESULT 6  
Q9ELT8 PRELIMINARY; PRT; 314 AA.  
AC Q9ELT8;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE REPLICATION ASSOCIATION PROTEIN.  
GN AC1.  
OS sweet potato leaf curl virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=100755;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
RT "Detection of a geminivirus infecting sweet potato in the United States."  
RL Plant Dis. 82:1253-1257(1998).  
RN [2]

RP SEQUENCE FROM N.A.  
RA Lotrakul P., Valverde R.A.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF288227; AAG01006.1; .  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
SQ SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 75.9%; Score 274; DB 12; Length 314;  
Best Local Similarity 79.4%; Pred. No. 3e-24; Mismatches 11; Indels 0; Gaps 0;  
Matches 54; Conservative 3;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFPQHNLNSNL 60  
DB 110 TITWGEFQVDRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPKYLFPQHNLVSNL 169

QY 61 DRIFDKTP 68

DB 170 DRIFSPPP 177

RESULT 7  
Q9QS55 PRELIMINARY; PRT; 364 AA.  
AC Q9QS55;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE REPLICATION INITIATION PROTEIN AC1.

GN AC1.  
OS sweet potato leaf curl virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=100755;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
RT "Detection of a geminivirus infecting sweet potato in the United States."  
RL Plant Dis. 82:1253-1257(1998).  
DR EMBL; AF104036; AAD47173.1; .  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 75.9%; Score 274; DB 12; Length 364;  
Best Local Similarity 79.4%; Pred. No. 3.5e-24; Mismatches 11; Indels 0; Gaps 0;  
Matches 54; Conservative 3;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFPQHNLNSNL 60  
DB 110 TITWGEFQVDRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPKYLFPQHNLVSNL 169

QY 61 DRIFDKTP 68

DB 170 DRIFSPPP 177

RESULT 8  
Q98693 PRELIMINARY; PRT; 185 AA.  
AC Q98693;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE REP PROTEIN (FRAGMENT).  
GN AC1.  
OS sida golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51034;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA;  
RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
RL Plant Dis. 81:1251-1258(1997).  
DR EMBL; U67926; AAB97865.1; .  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
FT NON\_TER 185 185  
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5BE1 CRC64;

Query Match 75.1%; Score 271; DB 12; Length 185;  
Best Local Similarity 70.0%; Pred. No. 3.6e-24; Mismatches 12; Indels 0; Gaps 0;  
Matches 49; Conservative 12;

QY 1 TLVWGEFQVDRSARGCGQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFPQHNLNSNL 60  
DB 89 TIEWGVFQIDGRSARGGQQTANDAAAEALNAGSKEDALKIIREKLPKYLFPQHNLSSNI 148

QY 61 DRIFDKTPEP 70

DB 149 DRIFSKPPEP 158

RESULT 9  
P88975

Query Match	72.0%	Score 260;	DB 12;	Length 234;
Best Local Similarity	68.6%	Pred. No. 9.4e-23;		
Matches 48;	Conservative 12;	Mismatches 10;	Indels 0;	Gaps 0;
QY	1	TLVGEFQVDRSGRGCGQTSDNAAAEALNASSKEEALQIIAAAIPEKYLFOFHNLNSNL 60		
Db	110	TI EWGLFQIDGRSGRGCGQTANDAAAEALNASTGTEAKMKIIKEKLPKFLFOYHNLSSDL 169		
QY	61	DRIFDKTPEP 70		
Db	170	DRIFMKAPNP 179		
RESULT 12				
Q9Z089				
ID	Q9Z089	PRELIMINARY;	PRT;	190 AA.
AC	Q9Z089;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL AND COMPLETE CDS			
DE	(FRAGMENT).			
GN	C1.			
OS	tobacco leaf curl virus.			
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.			
OX	NCBI_TaxID=67762;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STRATN=YOKOHAMA3;			
RC	Gai K., Ohshita S., Ishii I., Yahara T.;			
RT	"Molecular phylogeny of geminivirus infecting wild plants in Japan.";			
FL	J. Plant Res. 110:247-257(1997).			
DR	EMBL; AB001315; BAA34033.1; -			
DR	InterPro; IPR001191; GeminI_AL1.			
DR	Pfam; PF00799; GeminI_AL1; 1..			

Search completed: January 3, 2002, 15:56:25

Job time: 1121 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:38:49 ; Search time 65.28 Seconds  
(without alignments)  
24.130 Million cell updates/sec

Title: US-09-289-346A-4  
Perfect score: 361  
Sequence: 1 TLVWGEFQVDRSARGCOT.....FOFHNLNSLDRIFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	61.5	359	4	US-08-809-103B-2
2	222	61.5	359	4	US-08-809-103B-4
3	222	61.5	359	4	US-08-809-103B-6
4	222	61.5	359	4	US-08-809-103B-8
5	219	60.7	353	4	US-08-838-151A-44
6	219	60.7	353	4	US-08-838-151A-46
7	219	60.7	353	4	US-08-838-151A-49
8	219	60.7	353	4	US-08-838-151A-52
9	219	60.7	353	4	US-08-838-151A-55
10	214	59.3	361	4	US-08-838-151A-2
11	214	59.3	361	4	US-08-838-151A-4
12	214	59.3	361	4	US-08-838-151A-6
13	214	59.3	361	4	US-08-838-151A-8
14	200	55.4	357	4	US-08-838-151A-24
15	200	55.4	357	4	US-08-838-151A-27
16	200	55.4	357	4	US-08-838-151A-30
17	192	53.2	357	4	US-08-838-151A-20
18	60.5	16.8	131	4	US-08-838-151A-14
19	60.5	16.8	131	4	US-08-838-151A-14
20	58.5	16.2	356	1	US-08-781-562-1
21	57	15.8	602	2	US-08-419-652-6
22	57	15.8	771	1	US-07-923-976-6
23	57	15.8	783	6	5422248-2
24	57	15.8	836	1	US-07-923-976-4
25	57	15.8	863	1	US-07-923-976-8
26	55	15.2	783	6	5231168-2
27	54.5	15.1	258	4	US-09-251-645-5

28	54.5	15.1	1713	3	US-08-600-982-24	Sequence 24, Appl
29	54.5	15.1	1713	5	PCT-US94-10261A-24	Sequence 24, Appl
30	54	15.0	271	1	US-08-276-919-10	Sequence 10, Appl
31	54	15.0	271	1	US-08-776-088-13	Sequence 13, Appl
32	54	15.0	271	5	PCT-US95-09145A-13	Sequence 13, Appl
33	54	15.0	274	1	US-08-776-088-15	Sequence 15, Appl
34	54	15.0	274	5	PCT-US95-09145A-15	Sequence 15, Appl
35	54	15.0	325	1	US-08-276-919-4	Sequence 4, Appl
36	54	15.0	325	1	US-08-276-919-13	Sequence 13, Appl
37	54	15.0	325	1	US-08-776-088-4	Sequence 4, Appl
38	54	15.0	325	1	US-08-776-088-18	Sequence 18, Appl
39	54	15.0	325	5	PCT-US95-09145A-4	Sequence 4, Appl
40	54	15.0	325	5	PCT-US95-09145A-18	Sequence 18, Appl
41	54	15.0	354	1	US-08-276-919-2	Sequence 2, Appl
42	54	15.0	354	1	US-08-776-088-2	Sequence 2, Appl
43	54	15.0	354	1	US-08-776-088-6	Sequence 6, Appl
44	54	15.0	354	3	US-09-325-320-2	Sequence 2, Appl
45	54	15.0	354	4	US-09-585-109-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
; Sequence 2, Application US/08809103B  
; Patent No. 6133505  
; GENERAL INFORMATION:  
; APPLICANT: GRONENBORN, Bruno  
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
; TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YOUNG & THOMPSON  
; STREET: 745 South 23rd Street  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/809,103B  
; FILING DATE: 17-MAR-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 94.111040  
; FILING DATE: 15-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/FR95/01192  
; FILING DATE: 15-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PATCH, Andrew J.  
; REGISTRATION NUMBER: 32,925  
; REFERENCE/DOCKET NUMBER: US94AL CNR TOM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 521-2297  
; TELEX: 248425 EMBON  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-809-103B-2

Query Match 61.5%; Score 222; DB 4; Length 359;  
Best Local Similarity 58.0%; Pred. No. 3.2e-22;



STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809.103B  
FILING DATE: 17-MAR-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 94.11040  
FILING DATE: 15-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR95/01192  
FILING DATE: 15-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: PATCH, Andrew J.  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 521-2237  
TELEFAX: (703) 685-0573  
TELEX: 248425 EMBON  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-809-103B-8

Query Match 61.5%; Score 222; DB 4; Length 359;  
Best Local Similarity 58.0%; Pred. No. 3.2e-22;  
Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 2 LVWGEFQVDRSGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLSNLD 61  
DB 111 LWGTFQIDGRSGCGQQTANDAYAKAINAGSKQALDVIKELAPRDYVLFHFNLSNLD 170  
QY 62 RIFDKTPEP 70  
DB 171 KVFQVPPAP 179

RESULT 5  
US-08-838-151A-44  
Sequence 44, Application US/08838151A  
Patent No. 6291743  
GENERAL INFORMATION:  
APPLICANT: Stout, John T  
APPLICANT: Luu, Hang T  
APPLICANT: Maxwell, Douglas  
APPLICANT: Ahlquist, Paul  
APPLICANT: Hanson, Steve  
TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
TITLE OF INVENTION: Genes  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
STREET: Two Prudential Plaza, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838.151A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0260  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5400  
TELEFAX: 312-616-5460  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-838-151A-44

Query Match 60.7%; Score 219; DB 4; Length 353;  
Best Local Similarity 60.0%; Pred. No. 8.1e-22;  
Matches 42; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGCGQTSNDAAAEALNASKEEALQIIAAAIPEKYLFOFHNLSNLD 60  
DB 110 TIEWGQFQVDRSGCGQGSANDSYAKALNADSIESTILKEQPKDYVLFQHHNIRSNL 169  
QY 61 DRIFDKTPEP 70  
DB 170 ERIFVKVPEP 179

RESULT 6  
US-08-838-151A-46  
Sequence 46, Application US/08838151A  
Patent No. 6291743  
GENERAL INFORMATION:  
APPLICANT: Stout, John T  
APPLICANT: Luu, Hang T  
APPLICANT: Maxwell, Douglas  
APPLICANT: Ahlquist, Paul  
APPLICANT: Hanson, Steve  
TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
STREET: Two Prudential Plaza, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838.151A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0260  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5400  
TELEFAX: 312-616-5460  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: amino acid



```

RESULT 10.
US-08-838-151A-2
; Sequence 2, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Genminivirus
; TITLE OF INVENTION: Genes
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,151A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Mueller, Lisa V
REGISTRATION NUMBER: 38,978
REFERENCE/DOCKET NUMBER: SVS3801P0260
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5400
TELEFAX: 312-616-5460
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

```









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:37:37 ; Search time 144.17 Seconds  
(without alignments)  
35.965 Million cell updates/sec

Title: US-09-289-346A-5

Perfect score: 359

Sequence: 1 TLVGFQVDFGRSARGCQT.....FQFHLNLSNLDRIEDKTPPE 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

```
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	359	100.0	70 21 AAB18681	Mutant peptide der
2	343	95.5	70 21 AAB18677	Peptide fragment f
3	343	95.5	356 21 AAB18687	Amino acid sequenc
4	335	93.3	70 21 AAB18685	Mutant peptide der
5	332	92.5	70 21 AAB18688	Mutant peptide der
6	331	92.2	70 21 AAB18692	Mutant peptide der
7	329	91.6	70 21 AAB18684	Mutant peptide der
8	329	91.6	70 21 AAB18690	Mutant peptide der
9	328	91.4	70 21 AAB18678	Mutant peptide der
10	328	91.4	70 21 AAB18686	Mutant peptide der
11	327	91.1	70 21 AAB18689	Mutant peptide der

12	325	90.5	70 21 AAB18680	Mutant peptide der
13	325	90.5	70 21 AAB18691	Mutant peptide der
14	321	89.4	70 21 AAB18683	Mutant peptide der
15	319	88.9	70 21 AAB18682	Mutant peptide der
16	311	86.6	70 21 AAB18679	Mutant peptide der
17	220	61.3	361 18 AAW34336	Tomato mottle viru
18	220	61.3	361 18 AAW34324	Tomato mottle viru
19	220	61.3	361 18 AAW34325	Tomato mottle viru
20	220	61.3	361 18 AAW34326	Tomato mottle viru
21	219	61.0	359 17 AAR88870	Sardinian tomato y
22	219	61.0	359 17 AAR88871	Sardinian tomato y
23	219	61.0	359 17 AAR88872	Sardinian tomato y
24	218	60.7	353 18 AAW34338	Bean golden mosaic
25	218	60.7	353 18 AAW34332	Bean golden mosaic
26	218	60.7	353 18 AAW34333	Bean golden mosaic
27	218	60.7	353 18 AAW34334	Bean golden mosaic
28	218	60.7	353 18 AAW34335	Bean golden mosaic
29	216	60.2	353 8 AAP70407	ORF 4 gene product
30	208.5	58.1	361 8 AAP70562	Product of ORF 4 f
31	208	57.9	362 19 AAW56495	Tobacco leaf curl
32	203	56.5	357 18 AAW34329	Tomato yellow leaf
33	203	56.5	357 18 AAW34330	Tomato yellow leaf
34	203	56.5	357 18 AAW34331	Tomato yellow leaf
35	195	54.3	357 18 AAW34337	Tomato yellow leaf
36	66.5	18.5	512 19 AAW68473	HIV-1 strain YBF30
37	63.5	17.7	1693 21 AAB48457	Human laminin 5 po
38	63.5	17.7	1693 21 AAB48459	Human laminin 5 po
39	63.5	17.7	1713 16 AAR70148	Deduced sequence o
40	63.5	17.7	1713 21 AAB48458	Human laminin 5 po
41	63.5	17.7	1724 21 AAB48456	Human laminin 5 po
42	60.5	16.9	131 18 AAW34327	Tomato mottle viru
43	60.5	16.9	447 21 AAY93659	Human GTPase assoc
44	60.5	16.9	447 21 AAY87089	Human secreted pro
45	60.5	16.9	447 22 AAB65267	Human PRO1125 (UNQ

#### ALIGNMENTS

```
RESULT 1
AAB18681
ID AAB18681 standard; peptide; 70 AA.
XX
AC AAB18681;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (All) protein.
XX
KW Geminivirus; replication protein; Rep protein; All; transgenic plant;
KW Ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 47
FT FT /note= "wild type residue replaced with Ala"
FT Misc-difference 48
FT FT /note= "wild type residue replaced with Ala"
FT Misc-difference 49
FT FT /note= "wild type residue replaced with Ala"
```

WO200054573-A1.

21-SEP-2000.

15-MAR-2000; 2000WO-US06759.

18-MAR-1999; 99US-0125004.

09-APR-1999; 99US-0289346.

(UYN-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 XX Claim 52; Page 44; 73pp; English.  
 XX  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 XX Sequence 70 AA;  
 SQ

Query Match 100.0%; Score 359; DB 21; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-39;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRGARGCGCOTSDAAAEALNASSKEEALQIIREKIPAAALFQPHNLNSNL 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1 tlwgefvdgrsargcgctsdnaaaalnasskeaalgiirekipaaalfqphnlinsnl 60

QY 61 DRIFDKTPEP 70  
 Db |||||||||||  
 61 drifdktppep 70

RESULT 2  
 AAB18677  
 ID AAB18677 standard; peptide; 70 AA.  
 AC AAB18677;  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 XX Disclosure; Page 18; 73pp; English.  
 XX  
 XX The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX  
 XX Sequence 70 AA;  
 SQ

Query Match 95.5%; Score 343; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 3e-37;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRGARGCGCOTSDAAAEALNASSKEEALQIIREKIPAAALFQPHNLNSNL 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1 tlwgefvdgrsargcgctsdnaaaalnasskeaalgiirekipkylfqphnlinsnl 60

QY 61 DRIFDKTPEP 70  
 Db |||||||||||  
 61 drifdktppep 70

RESULT 3  
 AAB18687  
 ID AAB18687 standard; peptide; 356 AA.  
 AC AAB18687;  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Amino acid sequence of a geminivirus replication protein of TGMV.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 354  
 FT /note= "unspecified amino acid"  
 XX  
 PN WO200054573-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX  
 PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
XX  
PS Disclosure: Page 47-48; 73pp; English.  
XX  
XX The present sequence represents a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
XX  
SQ Sequence 356 AA;

Query Match 95.5%; Score 343; DB 21; Length 356;  
Best Local Similarity 95.7%; Pred. No. 2.6e-36;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TLVWGFQVGRSARGCGTSDNDAAEALNASSKEBALQIIREKIPAAALFQFHNLNSNL 60  
Db 110 TLVWGFQVGRSARGCGTSDNDAAEALNASSKEBALQIIREKIPAAALFQFHNLNSNL 169  
OY 61 DRIFDKTPEP 70  
Db 170 drifdktp 179

RESULT 4  
AAB18685  
ID AAB18685 standard; peptide; 70 AA.  
XX  
AC AAB18685;  
XX  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 10 /note= "wild type residue replaced with Ala"  
XX  
XX WO200054573-A1.  
XX  
PD 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNCL) UNIV NORTH CAROLINA STATE.  
XX  
PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX WPI; 2000-618851/59.  
XX  
PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
XX  
PS Claim 53; Page 46; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA;

Query Match 93.3%; Score 335; DB 21; Length 70;  
Best Local Similarity 94.3%; Pred. No. 3.3e-36;  
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TLVWGFQVGRSARGCGTSDNDAAEALNASSKEBALQIIREKIPAAALFQFHNLNSNL 60  
Db 1 TLVWGFQVGRSARGCGTSDNDAAEALNASSKEBALQIIREKIPAAALFQFHNLNSNL 60  
OY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70

RESULT 5  
AAB18688  
ID AAB18688 standard; peptide; 70 AA.  
XX  
AC AAB18688;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 19 /note= "wild type residue replaced with Ala"  
XX  
XX WO200054573-A1.  
XX  
PD 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNCL) UNIV NORTH CAROLINA STATE.  
XX  
PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX  
XX WPI; 2000-618851/59.  
XX  
DR

XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the RB binding region  
 PT  
 XX  
 PS Disclosure; Page 48; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;

Query Match 92.5%; Score 332; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 8.2e-36;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLVWGFEQVDGRSGCGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQPHNLSNL 60  
 Db 1 TLVWGFEQVDGRSGCGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQPHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 6  
 AAB18692 ID AAB18692 standard; peptide; 70 AA.  
 AC AAB18692;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 66 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 69 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 XX  
 XX 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 XX  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNV-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI

XX WPI; 2000-618851/59.  
 DR Transgenic plants with increased resistance to geminivirus infection  
 XX comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the RB binding region  
 PT  
 XX  
 PS Disclosure; Page 50; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 Query Match 92.2%; Score 331; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 1.1e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 TLVWGFEQVDGRSGCGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQPHNLSNL 60  
 Db 1 TLVWGFEQVDGRSGCGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQPHNLSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70

RESULT 7  
 AAB18684 ID AAB18684 standard; peptide; 70 AA.  
 AC AAB18684;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 7 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 8 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 XX  
 XX 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNV-) UNIV NORTH CAROLINA STATE.  
 PA

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant ALI protein with a mutation in the Rb binding region  
 PT -  
 XX Claim 52; Page 45; 73pp; English.  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as ALI. ALI binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the ALI  
 CC protein are used to produce transgenic plants. The mutation in ALI is  
 CC present in a ribosome binding region, and expression of mutant ALI  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant ALI proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX Sequence 70 AA:  
 SQ

Query Match 91.6%; Score 329; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 2e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDCGRSGCQTSDAAAEALNASSKEEALQIREKIPAAALFOFHNLNSNL 60  
 Db 1 TLVWGEFQVDCGRSGCQTSDAAAEALNASSKEEALQIREKIPAAALFOFHNLNSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktppep 70

RESULT 8  
 AAB18690  
 ID AAB18690 standard; peptide: 70 AA.  
 XX AAB18690;  
 AC AAB18690;  
 XX 22-JAN-2001 (first entry)  
 DT Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.  
 DE Geminivirus; replication protein; Rep protein; ALI; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX Synthetic.  
 OS Tomato golden mosaic virus.  
 XX Key Location/Qualifiers  
 FH Misc-difference 27  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 30  
 FT /note= "wild type residue replaced with Ala"  
 FT /note= "wild type residue replaced with Ala"  
 XX WO200054573-A1.  
 PN 21-SEP-2000.  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF 18-MAR-1999; 99US-0125004.  
 XX 09-APR-1999; 99US-0289346.  
 PR

XX (UVNC-) UNIV NORTH CAROLINA STATE.  
 PA Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI: 2000-618851/59.  
 DR Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant ALI protein with a mutation in the Rb binding region  
 PT -  
 XX Disclosure; Page 49; 73pp; English.  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as ALI. ALI binds  
 CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the ALI  
 CC protein are used to produce transgenic plants. The mutation in ALI is  
 CC present in a ribosome binding region, and expression of mutant ALI  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant ALI proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX Sequence 70 AA:  
 SQ

Query Match 91.6%; Score 329; DB 21; Length 70;  
 Best Local Similarity 92.9%; Pred. No. 2e-35;  
 Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDCGRSGCQTSDAAAEALNASSKEEALQIREKIPAAALFOFHNLNSNL 60  
 Db 1 TLVWGEFQVDCGRSGCQTSDAAAEALNASSKEEALQIREKIPAAALFOFHNLNSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktppep 70

RESULT 9  
 AAB18678  
 ID AAB18678 standard; peptide: 70 AA.  
 XX AAB18678;  
 AC AAB18678;  
 XX 22-JAN-2001 (first entry)  
 DT Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.  
 DE Geminivirus; replication protein; Rep protein; ALI; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX Synthetic.  
 OS Tomato golden mosaic virus.  
 XX Key Location/Qualifiers  
 FH Misc-difference 12  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 13  
 FT /note= "wild type residue replaced with Ala"  
 FT Misc-difference 15  
 FT /note= "wild type residue replaced with Ala"  
 XX WO200054573-A1.  
 PN 21-SEP-2000.  
 XX



```
FT      /note= "wild type residue replaced with Ala"  
PN      WO200054573-A1.  
XX      21-SEP-2000.  
XX      15-MAR-2000; 2000WO-US06759.  
XX      18-MAR-1999; 99US-0125004.  
PR      09-APR-1999; 99US-0289346.  
XX      (UYNC-) UNIV NORTH CAROLINA STATE.  
XX      Hanley-Bowdoin L, Orozco BM, Kong L;  
DR      WPI; 2000-618851/59.  
XX      Transgenic plants with increased resistance to geminivirus infection  
PT      comprise a nucleic acid construct containing a nucleic acid sequence  
PT      encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT      -  
XX      Disclosure; Page 48-49; 73pp; English.  
XX      The present sequence represents a mutant peptide, derived from a  
CC      geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC      double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC      DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC      protein are used to produce transgenic plants. The mutation in AL1 is  
CC      present in a ribosome binding region, and expression of mutant AL1  
CC      protein imparts increased resistance to geminivirus infection in the  
CC      plant. Mutant AL1 proteins are useful for producing plants having  
CC      increased resistance or reduced sensitivity to a geminivirus such as  
CC      tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC      virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC      cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC      virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC      virus, cotton leaf curl virus or beet curly top virus.  
XX      Sequence 70 AA;  
SQ  
  
Query Match 91.1%; Score 327; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 3.7e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 TLVWGEFQVGRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 60  
Db 1 tlwvgefvdgrsargcgctsaataaaalnasskeaalqiiirekipekylfqfhnlnsl 60  
  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70  
  
RESULT 12  
AAB18680  
ID AAB18680 standard; peptide; 70 AA.  
XX AAB18680;  
XX  
XX 22-JAN-2001 (first entry)  
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 42
```

```
FT      /note= "wild type residue replaced with Ala"  
FT      Misc-difference 43  
FT      /note= "wild type residue replaced with Ala"  
FT      Misc-difference 44  
FT      /note= "wild type residue replaced with Ala"  
XX  
PN      WO200054573-A1.  
XX      21-SEP-2000.  
XX      15-MAR-2000; 2000WO-US06759.  
XX      18-MAR-1999; 99US-0125004.  
PR      09-APR-1999; 99US-0289346.  
XX      (UYNC-) UNIV NORTH CAROLINA STATE.  
XX      Hanley-Bowdoin L, Orozco BM, Kong L;  
DR      WPI; 2000-618851/59.  
XX      Transgenic plants with increased resistance to geminivirus infection  
PT      comprise a nucleic acid construct containing a nucleic acid sequence  
PT      encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT      -  
XX      Claim 52; Page 43-44; 73pp; English.  
XX      The present sequence represents a mutant peptide, derived from a  
CC      geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC      double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC      DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC      protein are used to produce transgenic plants. The mutation in AL1 is  
CC      present in a ribosome binding region, and expression of mutant AL1  
CC      protein imparts increased resistance to geminivirus infection in the  
CC      plant. Mutant AL1 proteins are useful for producing plants having  
CC      increased resistance or reduced sensitivity to a geminivirus such as  
CC      tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC      virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC      cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC      virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC      virus, cotton leaf curl virus or beet curly top virus.  
XX      Sequence 70 AA;  
SQ  
  
Query Match 90.5%; Score 325; DB 21; Length 70;  
Best Local Similarity 91.4%; Pred. No. 6.7e-35;  
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
QY 1 TLVWGEFQVGRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 60  
Db 1 tlwvgefvdgrsargcgctsndaaalnasskeaalqliaaapekylfqfhnlnsl 60  
  
QY 61 DRIFDKTPEP 70  
Db 61 drifdktp 70  
  
RESULT 13  
AAB18691  
ID AAB18691 standard; peptide; 70 AA.  
XX AAB18691;  
XX  
XX 22-JAN-2001 (first entry)  
XX Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Synthetic.
```





Job time: 154 sec

DT  
XX  
DE  
XX  
KW  
KW  
XX  
OS  
OS  
XX  
PH  
FT  
FT  
FT  
FT  
FT  
XX  
PN  
XX  
PD  
XX  
PF  
XX  
PR  
PR  
XX  
PA  
XX  
PI  
XX  
DR  
XX  
PT  
PT  
PT  
XX  
PS  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ

22-JAN-2001 (first entry)  
Mutant peptide derived from amino acids 110-179 of Rep (All) protein.  
Geminivirus; replication protein; Rep protein; All; transgenic plant;  
ribosome binding region; resistance; geminivirus infection.  
Synthetic.  
Tomato golden mosaic virus.  
Key Location/Qualifiers  
Misc-difference 52 /note= "wild type residue replaced with Ala"  
Misc-difference 54 /note= "wild type residue replaced with Ala"  
Misc-difference 55 /note= "wild type residue replaced with Ala"  
WO200054573-Al.  
21-SEP-2000.  
15-MAR-2000; 2000WO-US06759.  
18-MAR-1999; 99US-0125004.  
09-APR-1999; 99US-0289346.  
(UYNC-) UNIV NORTH CAROLINA STATE.  
Hanley-Bowdoin L, Orozco BM, Kong L;  
WPI; 2000-618851/59.  
Transgenic plants with increased resistance to geminivirus infection  
comprise a nucleic acid construct containing a nucleic acid sequence  
encoding a mutant All protein with a mutation in the Rb binding region  
-  
Claim 53; Page 44-45; 73pp; English.  
The present sequence represents a mutant peptide, derived from a  
geminivirus replication (Rep) protein, also known as All. All binds  
double-stranded DNA, catalyses cleavage and ligation of single-stranded  
DNA, and interacts with other viral and host proteins. Mutants of the All  
protein are used to produce transgenic plants. The mutation in All is  
present in a ribosome binding region, and expression of mutant All  
protein imparts increased resistance to geminivirus infection in the  
plant. Mutant All proteins are useful for producing plants having  
increased resistance or reduced sensitivity to a geminivirus such as  
tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
virus, cotton leaf curl virus or beet curly top virus.

Query Match 88.9%; Score 319; DB 21; Length 70;  
Best Local Similarity 91.4%; Pred. NO. 4.le-34;  
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Oy 1 TLVWGEFQVDRSGGOTSNDAAAEALNASSKEEALQILREKIPAAALFOFHNLNSNL 60  
Dy ||||||||||||||||||||||||||||||||||||||||||||| || |||||  
1 tlwgefqvdrsggqtsndaaaealnasskeeaqlilrekipekyifafaalnsnl 60  
Oy 61 DRIFDKTPEP 70  
Dy |||||||||  
61 drifdktp 70



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:40:09 ; Search time 72.79 Seconds  
(without alignments)

Title: US-09-289-346A-5  
perfect score: 359  
Sequence: 1 TLVWGEFQVDGRSARGGCOT ..... FOFHNLNSNLDRIEDKTPPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database :      PIR_68:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	343	95.5	352	1	QOCVLI	ALL protein - toma	
2	247	68.8	361	1	QOCVPT	ALL protein - pota	
3	243	67.7	358	2	S07594	hypothetical prote	
4	232	64.6	362	1	QJ1887	ALL protein - toma	
5	226	63.0	359	2	S39211	gene C1 protein -	
6	224	62.4	349	2	QJ2305	replicase - pepper	
7	224	62.4	349	2	S31875	ALL protein - pepp	
8	219	61.0	359	2	S22593	hypothetical prote	
9	216	60.2	351	2	QJ2327	ALL protein - Indi	
10	216	60.2	358	1	QJ1870	ALL protein - toma	
11	215	59.9	355	1	QOCVMI	AVI protein - abut	
12	213	59.3	359	2	S39235	gene C1 protein -	
13	212	59.1	385	2	S28360	ALL protein - beet	
14	209	58.2	360	2	S59885	replication-associ	
15	203	56.5	357	1	QOCVCI	ALL protein - toma	
16	134	37.3	131	2	S45059	ACL protein (clone	
17	118	32.9	347	1	QOCVSI	ALL protein - squa	
18	65	18.1	587	2	JC1419	Fc gamma (IgG) rec	
19	63.5	17.7	1713	2	A55347	adhesive ligand ep	
20	62	17.3	160	2	G82060	hypothetical prote	
21	62	17.3	308	2	E86840	protein maturation	
22	61	17.0	631	2	S36505	E1 protein - human	
23	61	17.0	840	2	T36175	probable large ATP	
24	60.5	16.9	429	2	A44384	GTP-binding regula	
25	60.5	16.9	435	2	E59096	hypothetical prote	
26	60.5	16.9	447	2	T12544	hypothetical prote	
27	60.5	16.9	1070	1	A34600	1-phosphatidylinos	
28	60	16.7	397	2	B71078	probable NADH oxid	
29	59.5	16.6	706	2	H71707	hypothetical prote	

## ALIGNMENTS

```

RESULT      1
QCCVLL1
ALL protein - tomato golden mosaic virus
C:Species: tomato golden mosaic virus
A:Note: host Nicotiana sp. (tobacco)
C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C:Accession: A04170
R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of tomato
A:Reference number: A04163
A:Accession: A04170
A:Molecule type: DNA
A:Residues: 1-352 <HAM>
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C:Genetics:
A:Map position: segment A
C:Superfamily: tomato golden mosaic virus ALL protein

```

```
Query Match      68.8%; Score 247; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 1.5e-20;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
   |||||
Db 110 TIEMGLFQIDGRSARGGQQTVDNAAEALNSGPKKAAAMKIKKELPEKFLFQYHNLSNL 169

QY 61 DRIFDKTPE 69
   |||||
Db 170 DRIFMKAPE 178

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C:Species: cassava latent virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: S07594
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A:Reference number: S07590; MUID:90174930
A:Accession: S07594
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <NOR>
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C:Genetics:
A:Map position: segment DNAL
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      67.7%; Score 243; DB 2; Length 358;
Best Local Similarity 64.3%; Pred. No. 4.4e-20;
Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
   |||||
Db 109 TVEWGQFQIDGRSARGGQGSANDAYAKALNSGSKSEALNVIRELPKDFVLQFHNLSNL 168

QY 61 DRIFDKTPEP 70
   |||||
Db 169 DRIFQEPAP 178

RESULT 4
J01887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N:Alternate names: CL protein
C:Species: tomato yellow leaf curl virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C:Accession: J01887
R:Dry, I.B.; Riden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A:Reference number: J01885; MUID:93139778
A:Accession: J01887
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-362 <DRY>
A:Cross-references: GB:S53251
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      64.6%; Score 232; DB 1; Length 362;
Best Local Similarity 55.3%; Pred. No. 7.9e-19;
Matches 47; Conservative 8; Mismatches 14; Indels 16; Gaps 1;

QY 1 TLVWGEFQVDRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
   |||||
Db 110 TIEMGLFQIDGRSARGGQQTVDNAAEALNSGPKKAAAMKIKKELPEKFLFQYHNLSNL 169

Query Match      62.4%; Score 224; DB 2; Length 349;
Best Local Similarity 61.4%; Pred. No. 6.2e-18;
Matches 43; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
   |||||
Db 110 TVEWGQFQIDGRSARGGQGSANDTYAKALNSASAEALQIREEQPHFFLQFHNIVSNA 169

QY 61 DRIFDKTPEP 70
   |||||
Db 170 NRIFQTPPEP 179

RESULT 5
S39211
gene Cl protein - tomato yellow leaf curl virus
C:Species: tomato yellow leaf curl virus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C:Accession: S39211
R:Norris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
Submitted to the EMBL Data Library, August 1993
A:Description: High similarity among the tomato yellow leaf curl virus isolates from
A:Reference number: S39209
A:Accession: S39211
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <NOR>
A:Cross-references: EMBL:225751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      63.0%; Score 226; DB 2; Length 359;
Best Local Similarity 60.9%; Pred. No. 3.8e-18;
Matches 42; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVWGEFQVDRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 61
   |||||
Db 111 LVWGFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPDYILHFNHNSNLD 170

QY 62 RIFDKTPEP 70
   |||||
Db 171 RVFQVPPAP 179

RESULT 6
J02300
replicase - pepper huasteco virus (component A)
N:Alternate names: ORF AL1 protein
C:Species: pepper huasteco virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C:Accession: J02300
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A:Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A:Reference number: J02299; MUID:94015007
A:Accession: J02300
A:Molecule type: DNA
A:Residues: 1-349 <TOR>
A:Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match      62.4%; Score 224; DB 2; Length 349;
Best Local Similarity 61.4%; Pred. No. 6.2e-18;
Matches 43; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
   |||||
Db 110 TVEWGQFQIDGRSARGGQGSANDTYAKALNSASAEALQIREEQPHFFLQFHNIVSNA 169

QY 61 DRIFDKTPEP 70
   |||||
Db 170 NRIFQTPPEP 179

RESULT 7
S31875
AL1 protein - pepper rizado amarillo virus
C:Species: pepper rizado amarillo virus
```

C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Sep-1999  
C:Accession: S31875  
R:Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu  
submitted to the EMBL Data Library, February 1993  
A:Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa  
A:Reference number: S31872  
A:Accession: S31875  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-349 <TOR>  
A:Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025  
A:Note: the source is designated as pepper huasteco virus  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 62.4%; Score 224; DB 2; Length 349;  
Best Local Similarity 61.4%; Pred. No. 6.2e-18;  
Matches 43; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 60  
Db 110 TVEWGEFQIDRSARGGQSANDYAKALNSASAEALQIIEEQPHFFLQFHNIVSNA 169  
Qy 61 DRIFDKTPEP 70  
Db 170 NRIFQTPPEP 179  
A:Molecule type: DNA

RESULT 8  
S22593  
hypothetical protein C4 - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Sep-1999  
C:Accession: S22593  
R:Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.  
Nucleic Acids Res. 19, 6763-6769, 1991  
A:Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartit  
A:Reference number: S22588; MUID:92107660  
A:Accession: S22593  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-359 <KHE>  
A:Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.0%; Score 219; DB 2; Length 359;  
Best Local Similarity 58.0%; Pred. No. 2.4e-17;  
Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

Qy 2 LVWGEFQVDRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 61  
Db 111 LEWGTFOIDRSARGGQQTANDAYAKAINAGSKSQALDIKELAPRDYVYLHFHNINSNL 170  
Qy 62 RIFDKTPEP 70  
Db 171 KVFQVPPAP 179  
A:Molecule type: DNA

RESULT 9  
JQ2327  
AL1 protein - Indian cassava mosaic virus  
N:Alternate names: replication-associated protein  
C:Species: Indian cassava mosaic virus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2327; S35883  
R:Hong, Y.G.; Robinson, D.J.; Harrison, B.D.  
J. Gen. Virol. 74, 2437-2443, 1993  
A:Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran  
A:Reference number: JQ2326; MUID:94065670  
A:Accession: JQ2327  
A:Molecule type: DNA

A:Residues: 1-351 <HON>  
A:Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.2%; Score 216; DB 2; Length 351;  
Best Local Similarity 62.7%; Pred. No. 5.1e-17;  
Matches 42; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 4 WGEFQVDRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 63

Db 113 WGTFOIDRSARGGQSANDYAAALNSGSKSEALKILRELAPRDYLRDFHHITSSNL 172

Qy 64 FDKTPEP 70

Db 173 FTKPPPP 179

RESULT 10

JQ1870

AL1 protein - tomato mottle virus (isolate Florida)

C:Species: tomato mottle virus

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999

C:Accession: JQ1870

R:Abouzid, A.M.; Polston, J.E.; Hiebert, E.

J. Gen. Virol. 73, 3225-3229, 1992

A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f

A:Reference number: JQ1869; MUID:93107858

A:Accession: JQ1870

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-358 <ABO>

A:Cross-references: GB:L14460

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.2%; Score 216; DB 1; Length 358;  
Best Local Similarity 57.1%; Pred. No. 5.2e-17;  
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 60

Db 107 TIWGEFQIDRSARGGQSANDSYAKALNAGSVQSLAVLREEQPKDFVLQHNIRSNL 166

Qy 61 DRIFDKTPEP 70

Db 167 ERIFAKAPEP 176

RESULT 11

QOCVWL

AV1 protein - abutilon mosaic virus (isolate West India)

C:Species: abutilon mosaic virus

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Apr-1994

C:Accession: A36214

R:Frishmuth, T.; Zimmat, G.; Jeske, H.

Virol. 178, 461-468, 1990

A:Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as

A:Reference number: A36214; MUID:91020984

A:Accession: A36214

A:Molecule type: DNA

A:Residues: 1-355 <PRI>

A:Cross-references: EMBL:X15983

C:Genetics:

A:Map position: segment A

C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 59.9%; Score 215; DB 1; Length 355;  
Best Local Similarity 58.6%; Pred. No. 6.7e-17;  
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;





---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:57:15 ; Search time 43.68 seconds  
(without alignments)  
58.758 Million cell updates/sec

Title: US-09-289-346A-5

Perfect score: 359

Sequence: 1 TLVWGEFQVGRSARGCOT.....FOFHNLNSNLDRIKTPPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	343	95.5	352	1	VALL_TGMV
2	247	68.8	361	1	VALL_PYMV
3	243	67.7	358	1	VALL_CLVK
4	243	67.7	358	1	VALL_CLVN
5	232	64.6	362	1	VALL_TYLC
6	226	63.0	359	1	VALL_TYLCU
7	224	62.4	349	1	VALL_PHVU
8	219	61.0	359	1	VALL_TYLCM
9	216	60.2	353	1	VALL_BGMV
10	216	60.2	361	1	VALL_TMOV
11	215	59.9	355	1	VALL_ABMVW
12	212	59.1	358	1	VALL_BCTV
13	203	56.5	357	1	VALL_TYLCV
14	118	32.9	347	1	VALL_SLVC
15	65	18.1	630	1	VE1_HPV66
16	63.5	17.7	1713	1	LMA3_HUMAN
17	62.5	17.4	428	1	GBAL_CAVAN
18	61	17.0	631	1	VE1_HPV30
19	60.5	16.9	447	1	TBL2_HUMAN
20	60.5	16.9	1070	1	P11B_HUMAN
21	59.5	16.6	706	1	Y006_RICPR
22	59.5	16.6	1053	1	UBA3_WHEAT
23	59.5	16.6	1610	1	CCAD_MESAU
24	59.5	16.6	2161	1	CCAD_HUMAN
25	59.5	16.6	2203	1	CCAD_RAT
26	59	16.4	703	1	GYS2_RAT
27	59	16.4	1852	1	CCAS_CYPCA
28	58	16.2	387	1	Y4PF_RHISN
29	58	16.2	1127	1	Y855_TREPA
30	58	16.2	1608	1	HLXA_SERMA
31	57.5	16.0	367	1	LHX4_MOUSE
32	57.5	16.0	511	1	HUTH_VIBCH
33	57	15.9	355	1	CRTB_RHOSH
34	57	15.9	428	1	FLT1_MOUSE
35	57	15.9	432	1	PROA_DEIRA
36	57	15.9	524	1	MPPA_MOUSE
37	57	15.9	524	1	MPPA_RAT
38	57	15.9	874	1	SLAP_BACLI
39	56.5	15.7	1070	1	P11B_RAT
40	56	15.6	129	1	RK12_PORPU
41	56	15.6	329	1	TC3A_CAEEL
42	56	15.6	427	1	FLT1_HUMAN
43	56	15.6	428	1	FLT1_RAT
44	55.5	15.5	266	1	ETFB_MYCLE
45	55.5	15.5	266	1	ETFB_MYCTU

## RESULT 1

VALL\_TGMV  
ID VALL\_TGMV STANDARD: PRT: 352 AA.  
AC P03567;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE ALL PROTEIN.  
GN AC1.  
OS Tomato golden mosaic virus (TGMV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID:10831;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;  
RT "Complete nucleotide sequence of the infectious cloned DNA components  
of tomato golden mosaic virus: potential coding regions and regulatory  
sequences."  
RL EMBO J. 3:2197-2205(1984).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch)  
CC -----  
DR EMBL; K02029; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A04170; QOCVLL.  
DR InterPro; IPR001191; Gemini\_All.  
DR Pfam; PF00799; Gemini\_All; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLL.  
DR ProDom; P000736; Gemini\_All; 1.  
DR ATP-binding.  
KW NP\_BIND 223 230 ATP (POTENTIAL).  
FT SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;  
SQ

Query Match 95.5%; Score 343; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 2e-31;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TLVWGEFQVGRSARGCOTSDNAAAEALNASSKEALQIIRKIPAAALFQFHNLNSNL 60  
Db 111 TLVWGEFQVGRSARGCOTSDNAAAEALNASSKEALQIIRKIPAAALFQFHNLNSNL 170  
OY 61 DRIFDKTPPEP 70  
Db 171 DRIFDKTPPEP 180

## RESULT 2

VALL\_PYMV

```

ID VAL1_PYMV STANDARD; PRT; 361 AA.
AC P27258;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE AL1 PROTEIN.
OS Potato yellow mosaic virus (isolate Venezuela).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10828;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311403; PubMed=18566690;
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT "The nucleotide sequence of the infectious cloned DNA components of
  potato yellow mosaic virus.";
RL J. Gen. Virol. 72:1515-1520(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00940; BAA00782.1; -
DR PIR: J00364; QOCVPT.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVLL.
DR ProDom: PD000736; Gemini_AL1; 1.
DR ATP-binding.
KW NP-BIND
FT NP-BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match 68.8%; Score 247; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 1.4e-20;
Matches 47; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGRSARGGQTSNDAAAEALNSKKEALQIREKIPAAALFOFHLNSNL 60
   I: |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||: ||
DB 110 TIEMGLFQIDGRSARGGQTVNDAAAEALNSKKEALQIREKIPAAALFOFHLNSNL 169

QY 61 DRIFDKTPE 69
   ||||| ||
DB 170 DRIFDKAPE 178

RESULT 3
VAL1_CLVWK
AC P14982;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (40.4 KDA PROTEIN).
GN AC1.
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RX Stanley J., Gay M.R.;
RA "Nucleotide sequence of cassava latent virus DNA.";
RL Nature 301:260-262(1983).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02057; -; NOT_ANNOTATED_CDS.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVLL.
DR ProDom: PD000736; Gemini_AL1; 1.
DR ATP-binding.
KW NP-BIND
FT NP-BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 67.7%; Score 243; DB 1; Length 358;
Best Local Similarity 64.3%; Pred. No. 3.8e-20;
Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 1 TLVWGEFQVGRSARGGQTSNDAAAEALNSKKEALQIREKIPAAALFOFHLNSNL 60
   I: |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||: ||
DB 109 TIEMGLFQIDGRSARGGQTSNDAYAKALNSKKEALNVIRELPKDFVLQFHLNSNL 168

QY 61 DRIFDKTPE 70
   ||||| ||
DB 169 DRIFQEPAP 178

RESULT 4
VAL1_CLVKN
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN (40.4 KDA PROTEIN).
GN AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174930; PubMed=2308831;
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT "Nucleotide sequence of the infectious cloned DNA components of
  African cassava mosaic virus (Nigerian strain).";
RL Nucleic Acids Res. 18:197-198(1990).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17095; CAA34953.1; -
DR PIR: S07594; S07594.
DR InterPro: IPR001191; Gemini_AL1.
DR Pfam: PF00799; Gemini_AL1; 1.
DR PRINTS: PR00227; GEMCOATALL.
DR PRINTS: PR00228; GEMCOATCLVLL.
DR ProDom: PD000736; Gemini_AL1; 1.
DR ATP-binding.
KW NP-BIND
FT NP-BIND 220 227 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40435 MW; 1DB16B80CB2D5E2C CRC64;

Query Match 67.7%; Score 243; DB 1; Length 358;
Best Local Similarity 64.3%; Pred. No. 3.8e-20;
Matches 45; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

```

-----  
 CC EMBL; X70418; CAA49856.1; --  
 CC PIR; S31875; S31875.  
 CC PIR; JQ2300; JQ2300.  
 CC InterPro; IPR001191; Gemini\_ALL.  
 CC Pfam; PF00779; Gemini\_ALL; 1.  
 CC PRINTS; PR00227; GEMCOATVALL.  
 CC PRINTS; PR00228; GEMCOATCLVLL.  
 -----

ProDom; PD000736; Gemini\_ALL; 1.  
 KW ATP-binding. 221 228 ATP (BY SIMILARITY).  
 FT NP\_BIND  
 SQ SEQUENCE 349 AA: 39722 MW; D5F4E76CD56370F4 CRC64;

Query Match 62.4%; Score 224; DB 1; Length 349;  
 Best Local Similarity 61.4%; Pred. No. 5.1e-18;  
 Matches 43; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSARGCGTSDNAAAEALNASSKEEALQIIREKIPAAALFOFHNLNSNL 60  
 DB 110 TVWGEFQIDGSRGQGSANDTYAKALNASSKEEALQIIREKIPAAALFOFHNLNSNL 169

QY 61 DRIFDKTPEP 70  
 DB 170 NRIFQTPPEP 179

RESULT 8  
 VAL1\_TYLCM STANDARD; PRT; 359 AA.  
 ID VAL1\_TYLCM AC P27360;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN (C1 PROTEIN).  
 GN C1.  
 OS Tomato yellow leaf curl virus (strain Marmande) (TYLCV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Khey-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S., Gronenborn B.;  
 RT "Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus."; Nucleic Acids Res. 19:6763-6769(1991).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M10070; AAA46318.1;  
 DR InterPro; IPR001191; Gemini\_ALL.  
 DR Pfam; PF00799; Gemini\_ALL; 1.  
 DR PRINTS; PR00227; GEMCOATALL.  
 DR PRINTS; PR00228; GEMCOATCLVL.  
 DR ProDom; PD000736; Gemini\_ALL; 1.  
 KW ATP-binding.  
 FT NP\_BIND 220 227 ATP (POTENTIAL).  
 SQ SEQUENCE 359 AA: 40733 MW; 9717B4A07C93EFA7 CRC64;

Query Match 61.0%; Score 219; DB 1; Length 359;  
 Best Local Similarity 58.0%; Pred. No. 1.9e-17;  
 Matches 40; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 2 LVWGEFQVDSARGCGTSDNAAAEALNASSKEEALQIIREKIPAAALFOFHNLNSNL 61  
 DB 111 LEWGTQFIDGSRGQGSANDTYAKALNASSKEEALQIIREKIPAAALFOFHNLNSNL 170

QY 62 RIFDKTPEP 70  
 DB 171 KVFQVPPAP 179

RESULT 10  
 VAL1\_TMOV STANDARD; PRT; 361 AA.  
 ID VAL1\_TMOV AC Q06657;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE AL1 PROTEIN.  
 GN AL1.  
 OS Tomato mottle virus (isolate Florida) (TMOV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=36449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Abouzid A.M., Polston J.E., Hiebert E.;  
 RT "The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from tomatoes in Florida."; J. Gen. Virol. 73:3225-3229(1992).  
 CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.

```
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14460; AAC32414.1; -
DR PIR; JQ1870; JQ1870.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
DR ATP-binding.
KW NP_BIND 222 229 ATP (BY SIMILARITY).
SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEAC6950 CRC64;

Query Match 60.2%; Score 216; DB 1; Length 361;
Best Local Similarity 57.1%; Pred. No. 4.3e-17;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 60
   I: |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||
Db 110 TIEMGDFDIDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 169

QY 61 DRIFDKTPEP 70
   :||| | |||
Db 170 ERIFAKAPEP 179

RESULT 11
VALL_ABMVW STANDARD; PRT; 355 AA.
ID VALL_ABMVW
AC P21947;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN Aci.
OS Abutilon mosaic virus (isolate West India).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91020984; PubMed=2219703;
RA Frischmuth T., Zimmat G., Jeske H.;
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT as well as eukaryotic features.";
RL Virology 178:461-468(1990).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR PIR; A36214; QCCWV1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
DR ATP-binding.
KW NP_BIND 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228
SQ SEQUENCE 355 AA; 40257 MW; 16A2CA8A63251E95 CRC64;

Query Match 60.2%; Score 216; DB 1; Length 361;
Best Local Similarity 57.1%; Pred. No. 4.3e-17;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 60
   I: |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||
Db 110 TIEMGDFDIDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 169

QY 61 DRIFDKTPEP 70
   :||| | |||
Db 170 ERIFAKAPEP 179

RESULT 11
VALL_ABMVW STANDARD; PRT; 355 AA.
ID VALL_ABMVW
AC P21947;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN Aci.
OS Abutilon mosaic virus (isolate West India).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91020984; PubMed=2219703;
RA Frischmuth T., Zimmat G., Jeske H.;
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT as well as eukaryotic features.";
RL Virology 178:461-468(1990).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR PIR; A36214; QCCWV1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
DR ATP-binding.
KW NP_BIND 221 228 ATP (POTENTIAL).
FT NP_BIND 221 228
SQ SEQUENCE 355 AA; 40257 MW; 16A2CA8A63251E95 CRC64;

Query Match 59.1%; Score 212; DB 1; Length 358;
Best Local Similarity 55.7%; Pred. No. 1.2e-16;
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 60
   I: |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||
Db 110 TIEMGFEFDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 169

QY 61 DRIFDKTPEP 70
   :|| | ||: ||
Db 170 QKIFORPPDP 179

RESULT 13
VALL_TYLCV STANDARD; PRT; 357 AA.
ID VALL_TYLCV
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
```

```
Query Match 59.9%; Score 215; DB 1; Length 355;
Best Local Similarity 58.6%; Pred. No. 5.4e-17;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 60
   I: |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||
Db 110 TAEWGEFVDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 169

QY 61 DRIFDKTPEP 70
   :||| | |||
Db 170 ERIFAKAPEP 179

RESULT 12
VALL_BCTV STANDARD; PRT; 358 AA.
ID VALL_BCTV
AC P14991;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE ALL PROTEIN (40.8 KDA PROTEIN).
OC Beet curly top virus (BCTV).
OS Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX NCBI_TaxID=10840;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT "The nucleotide sequence of an infectious clone of the geminivirus
RT beet curly top virus.";
RL EMBO J. 5:1761-1767(1986).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 59.1%; Score 212; DB 1; Length 358;
Best Local Similarity 55.7%; Pred. No. 1.2e-16;
Matches 39; Conservative 15; Mismatches 16; Indels 0; Gaps 0;

QY 1 TLVWGEFVDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 60
   I: |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: ||
Db 110 TIEMGFEFDGSRARGGQTSNDAAAEALNASSKEALQIIRKIPAAALFQFHNLSNL 169

QY 61 DRIFDKTPEP 70
   :|| | ||: ||
Db 170 QKIFORPPDP 179

RESULT 13
VALL_TYLCV STANDARD; PRT; 357 AA.
ID VALL_TYLCV
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
```

```
DE AL1 PROTEIN (C1 PROTEIN).
GN C1.
OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024070; PubMed=1926771;
RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
with a single genomic component.";
RL Virology 185:151-161(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X15656; CAA33688.1; -.
DR PIR; D40779; QOCVC1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1.1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLL.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 219 226 ATP (POTENTIAL).
SQ SEQUENCE 357 AA; 40678 MW; 939AB68E1AB3B2A7 CRC64;

Query Match 56.5%; Score 203; DB 1; Length 357;
Best Local Similarity 65.6%; Pred. No. 1.2e-15;
Matches 40; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 4 MGFEVDGRSARGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFOFHNLSNLDRI 63
DB 111 FGVSQIDGRSARGGQGSANDAYAEALNSGSKSEALNILEKAPKDYILOFHNLSSNLDRI 170

QY 64 F 64
DB 171 F 171

RESULT 14
VALL_SLCV STANDARD; PRT; 347 AA.
AC P29048.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE AL1 PROTEIN.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082449; PubMed=1984668;
RA Lazarowitz S.G., Lazdins I.B.;
RT "Infectivity and complete nucleotide sequence of the cloned genomic
components of a bipartite squash leaf curl geminivirus with a broad
host range phenotype.";
RL Virology 180:58-69(1991).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M38183; AAC32410.1; ALT_INIT.
DR PIR; C36785; QOCV51.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1.1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVLL.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 218 225 ATP (POTENTIAL).
SQ SEQUENCE 347 AA; 39110 MW; AFDAEBDE122110E CRC64;

Query Match 32.9%; Score 118; DB 1; Length 347;
Best Local Similarity 37.9%; Pred. No. 4.6e-06;
Matches 25; Conservative 12; Mismatches 25; Indels 4; Gaps 1;

QY 5 GFEVDGRSARGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFOFHNLSNLDRI 64
DB 116 GQYKV5G-----GSKSNKDDVYHNAVGSAGEALDIKAGDPRTFVYHNLNLANVERLF 171

QY 65 DKTPEP 70
DB 172 QKPEP 177

RESULT 15
VEL_HPV66 STANDARD; PRT; 630 AA.
AC Q80957.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE REPLICATION PROTEIN E1.
GN E1.
OS Human papillomavirus type 66.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Papillomavirus.
OX NCBI_TaxID=37119;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U31794; AAA79501.1; -.
DR InterPro; IPR001177; Papillom_E1.
DR Pfam; PF00519; E1; 1.
DR Pfam; PF00524; E1_N; 1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
Nuclear protein.
FT NP_BIND 458 465 ATP (POTENTIAL).
SQ SEQUENCE 630 AA; 72065 MW; 22DDDA5934F7291B CRC64;

Query Match 18.1%; Score 65; DB 1; Length 630;
Best Local Similarity 25.5%; Pred. No. 8.5;
Matches 13; Conservative 14; Mismatches 24; Indels 0; Gaps 0;
```



---

**THIS PAGE BLANK (USPTO)**



Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	296	82.5	352	12	Q9E000	Q9E000 tomato rugo	
2	287	79.9	226	12	Q9WHF6	Q9WHF6 tomato mild	
3	277	77.2	361	12	Q67574	Q67574 bean golden	
4	275	76.6	225	12	Q9QDB1	Q9QDB1 cowpea gold	
5	275	76.6	226	12	Q09727	Q09727 leonurus mo	
6	271	75.5	185	12	Q98693	Q98693 sida golden	
7	271	75.5	314	12	Q9ELT8	Q9ELT8 sweet potato	
8	271	75.5	364	12	Q9QS55	Q9QS55 sweet potato	
9	264	73.5	149	12	P88975	P88975 macroptiliu	
10	264	73.5	190	12	Q9Z089	Q9Z089 tobacco lea	
11	264	73.5	190	12	Q9Z084	Q9Z084 tobacco lea	
12	264	73.5	233	12	Q9Y1A4	Q9Y1A4 macroptiliu	
13	263	73.3	190	12	Q9W827	Q9W827 tobacco lea	
14	263	73.3	208	12	Q9Z0C4	Q9Z0C4 tobacco lea	
15	261	72.7	208	12	Q9Z0C0	Q9Z0C0 tobacco lea	
16	261	72.7	208	12	Q9Z0B8	Q9Z0B8 tobacco lea	
17	259	72.1	208	12	Q9Z0B6	Q9Z0B6 tobacco lea	
18	258	71.9	203	12	Q9Z0B3	Q9Z0B3 tobacco lea	
19	258	71.9	234	12	Q39180	Q39180 potato yell	

```
Q9WHF6
ID Q9WHF6 PRELIMINARY; PRT; 226 AA.
AC Q9WHF6
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN96-HSKW;
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131071; AAD33471.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 79.9%; Score 287; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 2.7e-25;
Matches 54; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDCRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 60
Db 111 TLVWGEFQVDCRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNL 170

QY 61 DRIFDKTPEP 70
Db 171 DRIFAKAPEP 180

RESULT 3
ID Q67574 PRELIMINARY; PRT; 361 AA.
AC Q67574
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE REPLICATIVE PROTEIN.
GN AL1.
OS Bean golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10839;
RN [1]
RP SEQUENCE FROM N.A.
RC Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA Maxwell D.P., Russell D.R.;
RL Phycopathology 81:980-985(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA Morales F.J., Maxwell D.P.;
RL Plant Dis. 75:336-342(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M88686; AAA46312.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1.
SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 77.2%; Score 277; DB 12; Length 361;
Best Local Similarity 77.6%; Pred. No. 6.5e-24;
Matches 52; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 4 WGEFQVDCRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNLDRI 63
Db 113 WGEFQVDCRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNLDRI 172

QY 64 FDKTPEP 70
Db 173 FTKAPDP 179

RESULT 4
ID Q9QDB1 PRELIMINARY; PRT; 225 AA.
AC Q9QDB1
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS cowpea golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=69263;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGMV-BR;
RA Faria J.C.;
RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT Brazil.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188708; AAF06318.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1.
FT NON_TER 225
SQ SEQUENCE 225 AA; 25766 MW; 1089CB6BD8D15B5D CRC64;

Query Match 76.6%; Score 275; DB 12; Length 225;
Best Local Similarity 77.6%; Pred. No. 6.5e-24;
Matches 52; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 4 WGEFQVDCRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNLDRI 63
Db 113 WGEFQVDCRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLSNLDRI 172

QY 64 FDKTPEP 70
Db 173 FKKPEP 179

RESULT 5
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV-BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; U92532; AAB51157.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 76.6%; Score 275; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 6.5e-24;
Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGGCGTSDNAAAEALNASSKEALQIIREKIPAAALFQFHNLSNL 60
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111 TVEWGEFQVDRSGRGCGTQVNDAAAEALNAPDKRTALQIIRKLPKYLQFHNLSNL 170

QY 61 DRIFDKTPEP 70
   ||||| |||||
Db 171 DRIFAKAPEP 180

RESULT 6
ID Q98693 PRELIMINARY; PRT; 185 AA.
AC Q98693;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN AC1.
OS sida golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51034;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA;
RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL Plant Dis. 81:1251-1258(1997).
DR EMBL; U67926; AAB97865.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
FT NON_TER 185
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 75.5%; Score 271; DB 12; Length 185;
Best Local Similarity 71.4%; Pred. No. 1.5e-23;
Matches 50; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGGCGTSDNAAAEALNASSKEALQIIREKIPAAALFQFHNLSNL 60
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 TIWGVFQIDGRSARGGCGTQNDAAAEALNAGSKEDALKIIRKLPKYLQFHNLSNL 148

QY 61 DRIFDKTPEP 70
   ||||| |||||
Db 149 DRIFSKPPEP 158

RESULT 7
ID Q9ELT8 PRELIMINARY; PRT; 314 AA.
AC Q9ELT8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATION PROTEIN.
GN AC1.
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
```

```
OX NCBI_TaxID=100755;
RN [1]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT "Detection of a geminivirus infecting sweet potato in the United
   States.";
RL Plant Dis. 82:1253-1257(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF288227; AAG01006.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
SQ SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 75.5%; Score 271; DB 12; Length 314;
Best Local Similarity 79.4%; Pred. No. 2.7e-23;
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGGCGTSDNAAAEALNASSKEALQIIREKIPAAALFQFHNLSNL 60
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 TITWGEFQVDRSARGGCGTQNDAAAEALNAGSKAALQIIREKLPKYLQFHNLSNL 169

QY 61 DRIFDKTP 68
   ||||| |
Db 170 DRIFSPPP 177

RESULT 8
ID Q9QS55 PRELIMINARY; PRT; 364 AA.
AC Q9QS55;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE REPLICATION INITIATION PROTEIN AC1.
GN AC1.
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=100755;
RN [1]
RP SEQUENCE FROM N.A.
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT "Detection of a geminivirus infecting sweet potato in the United
   States.";
RL Plant Dis. 82:1253-1257(1998).
DR EMBL; AF104036; AAD47173.1; -.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Gemini_AL1; 1.
SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 75.5%; Score 271; DB 12; Length 364;
Best Local Similarity 79.4%; Pred. No. 3.2e-23;
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGGCGTSDNAAAEALNASSKEALQIIREKIPAAALFQFHNLSNL 60
   I: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 TITWGEFQVDRSARGGCGTQNDAAAEALNAGSKAALQIIREKLPKYLQFHNLSNL 169

QY 61 DRIFDKTP 68
   ||||| |
Db 170 DRIFSPPP 177

RESULT 9
ID Q9ELT8 PRELIMINARY; PRT; 314 AA.
AC Q9ELT8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATION PROTEIN.
GN AC1.
OS sweet potato leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
```

```

ID P88975; PRELIMINARY; PRT; 149 AA.
AC
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN AC1.
OS Macroptilium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICAN;
RA Roye M.E.;
RL Thesis (1996); Biochemistry, University of the West Indies, Jamaica.
DR EMBL; 0752278; AAB36919.1;
DR InterPro; IP001191; Geminini_AL1.
DR Pfam; PF00799; Geminini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Geminini_AL1; 1.
FT NON_TER 1
FT NON_TER 149
FT NON_TER 149
SQ SEQUENCE 149 AA; 16785 MW; E4CF5ED4C9CD508 CRC64;

Query Match 73.5%; Score 264; DB 12; Length 149;
Best Local Similarity 70.0%; Pred. No. 7.5e-23;
Matches 49; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

QY 1 TLWGEFQVDRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFOFHNLSNL 60
   |||||
Db 52 TIWGVFQIDGRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFOFHNLSNL 111
   |||||

QY 61 DRIFDKTPEP 70
   |||||
Db 112 DRIFMKDPEP 121

RESULT 10
Q92089 PRELIMINARY; PRT; 190 AA.
AC
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Cl.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-YOKOHAMA3;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RL "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
DR EMBL; AB001318; BAA34033.1;
DR InterPro; IP001191; Geminini_AL1.
DR Pfam; PF00799; Geminini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Geminini_AL1; 1.
FT NON_TER 1
FT NON_TER 190
FT NON_TER 190
SQ SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

Query Match 73.5%; Score 264; DB 12; Length 190;
Best Local Similarity 62.4%; Pred. No. 9.9e-23;
Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLWGEFQVDRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFOFHNLSNL 60
   |||||
Db 85 TLEWGTFOIDGRSARGCGQNDACAEALNASSKAEALAIIREKLPKDFIFOYHNLSNL 144
   |||||

QY 61 DRIFAPPLEVFCPTASSFDQVPE 69
   |||||
Db 145 DRIFAPPLEVFCPTASSFDQVPE 169

RESULT 12
Q9YL44 PRELIMINARY; PRT; 233 AA.
AC
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS Macroptilium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E.;
RL "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses from Jamaica.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E., McLaughlin W.A., Maxwell D.P.;
RL "Molecular characterization of two distinct geminiviruses infecting M.
```

```

Db 85 TLEWGTFOIDGRSARGCGQNDACAEALNASSKAEALAIIREKLPKDFIFOYHNLSNL 144
   |||||
QY 61 DRIFAPPLEVFCPTASSFDQVPE 69
   |||||
Db 145 DRIFAPPLEVFCPTASSFDQVPE 169

RESULT 11
Q92084 PRELIMINARY; PRT; 190 AA.
AC
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL AND COMPLETE CDS (FRAGMENT).
GN Cl.
OS tobacco leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=67762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-YOKOHAMA5;
RA Ooi K., Ohshita S., Ishii I., Yahara T.;
RL "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
DR EMBL; AB001318; BAA34033.1;
DR InterPro; IP001191; Geminini_AL1.
DR Pfam; PF00799; Geminini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR ProDom; PD000736; Geminini_AL1; 1.
FT NON_TER 1
FT NON_TER 190
FT NON_TER 190
SQ SEQUENCE 190 AA; 21444 MW; AAC1C2943E3F01AD CRC64;

Query Match 73.5%; Score 264; DB 12; Length 190;
Best Local Similarity 62.4%; Pred. No. 9.9e-23;
Matches 53; Conservative 6; Mismatches 10; Indels 16; Gaps 1;

QY 1 TLWGEFQVDRSARGCGQTSNDAAAEALNASSKEEALQIIREKIPAAALFOFHNLSNL 60
   |||||
Db 85 TLEWGTFOIDGRSARGCGQNDACAEALNASSKAEALAIIREKLPKDFIFOYHNLSNL 144
   |||||

QY 61 DRIFAPPLEVFCPTASSFDQVPE 69
   |||||
Db 145 DRIFAPPLEVFCPTASSFDQVPE 169

RESULT 12
Q9YL44 PRELIMINARY; PRT; 233 AA.
AC
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS Macroptilium golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=51676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E.;
RL "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses from Jamaica.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-JAMAICA STRAIN 1;
RA Roye M.E., McLaughlin W.A., Maxwell D.P.;
RL "Molecular characterization of two distinct geminiviruses infecting M.
```



Search completed: January 3, 2002, 15:56:25  
Job time: 1121 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:38:49 ; Search time 65.28 Seconds  
(without alignments)  
24.130 Million cell updates/sec

Title: US-09-289-346A-5

Perfect score: 359

Sequence: 1 TLVWGEFQVDRSGRGCGQT.....FQFHNLNSLDRFDKTPPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PTDUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	220	61.3	361	US-08-838-151A-2	Sequence 2, Appl
2	220	61.3	361	US-08-838-151A-4	Sequence 4, Appl
3	220	61.3	361	US-08-838-151A-6	Sequence 6, Appl
4	220	61.3	361	US-08-838-151A-8	Sequence 8, Appl
5	219	61.0	359	US-08-809-103B-2	Sequence 2, Appl
6	219	61.0	359	US-08-809-103B-4	Sequence 4, Appl
7	219	61.0	359	US-08-809-103B-6	Sequence 6, Appl
8	219	61.0	359	US-08-809-103B-8	Sequence 8, Appl
9	218	60.7	353	US-08-838-151A-44	Sequence 44, Appl
10	218	60.7	353	US-08-838-151A-46	Sequence 46, Appl
11	218	60.7	353	US-08-838-151A-49	Sequence 49, Appl
12	218	60.7	353	US-08-838-151A-52	Sequence 52, Appl
13	218	60.7	353	US-08-838-151A-55	Sequence 55, Appl
14	203	56.5	357	US-08-838-151A-24	Sequence 24, Appl
15	203	56.5	357	US-08-838-151A-27	Sequence 27, Appl
16	203	56.5	357	US-08-838-151A-30	Sequence 30, Appl
17	195	54.3	357	US-08-838-151A-20	Sequence 20, Appl
18	63.5	17.7	1713	US-08-600-982-24	Sequence 24, Appl
19	63.5	17.7	1713	US-08-600-982-24	Sequence 24, Appl
20	60.5	16.9	131	US-08-838-151A-14	Sequence 14, Appl
21	59.5	16.6	2161	US-07-745-206A-2	Sequence 2, Appl
22	59.5	16.6	2161	US-08-455-543A-49	Sequence 49, Appl
23	59.5	16.6	2161	US-08-455-543A-51	Sequence 51, Appl
24	59.5	16.6	2161	US-08-223-305C-49	Sequence 49, Appl
25	59.5	16.6	2161	US-08-223-305C-51	Sequence 51, Appl
26	59.5	16.6	2161	US-08-311-363-2	Sequence 2, Appl
27	58.5	16.3	374	US-08-928-692-51	Sequence 51, Appl

28	56	15.6	171	2	US-08-609-049A-22	Sequence 22, Appl
29	56	15.6	171	4	US-09-170-996-22	Sequence 22, Appl
30	56	15.6	416	4	US-09-416-050A-4	Sequence 4, Appl
31	56	15.6	416	4	US-09-664-800-4	Sequence 4, Appl
32	56	15.6	416	4	US-09-665-309-4	Sequence 4, Appl
33	56	15.6	416	4	US-09-661-569-4	Sequence 4, Appl
34	54.5	15.2	3898	2	US-08-876-991-2	Sequence 2, Appl
35	54.5	15.2	3898	2	US-09-059-853-2	Sequence 2, Appl
36	54	15.0	242	3	US-08-622-352A-3	Sequence 3, Appl
37	54	15.0	242	3	US-08-826-390-3	Sequence 3, Appl
38	54	15.0	271	1	US-08-276-919-10	Sequence 10, Appl
39	54	15.0	271	1	US-08-776-088-13	Sequence 13, Appl
40	54	15.0	271	5	PCT-US95-09145A-13	Sequence 13, Appl
41	54	15.0	274	1	US-08-776-088-15	Sequence 15, Appl
42	54	15.0	274	5	PCT-US95-09145A-15	Sequence 15, Appl
43	54	15.0	325	1	US-08-276-919-4	Sequence 4, Appl
44	54	15.0	325	1	US-08-276-919-13	Sequence 13, Appl
45	54	15.0	325	1	US-08-776-088-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-08-838-151A-2  
; Sequence 2, Application US/08838151A  
; Patent No. 6251743  
; GENERAL INFORMATION:  
; APPLICANT: Stout, John T  
; APPLICANT: Luu, Hang T  
; APPLICANT: Maxwell, Douglas  
; APPLICANT: Ahlquist, Paul  
; APPLICANT: Hanson, Steve  
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
; TITLE OF INVENTION: Genes  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rockey, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,151A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mueller, Lisa V  
; REGISTRATION NUMBER: 38,978  
; REFERENCE/DOCKET NUMBER: SVS3801P0260  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-838-151A-2

Query Match 61.3%; Score 220; DB 4; Length 361;  
Best Local Similarity 58.6%; Pred. No. 2.3e-21;  
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TLVWGEFQVDRSGRGCGQTSDAAAFALNASSKEEALQIIREKIPAAALFQFHNLNSL 60

```

      | : ||||| ||||| : ||||| : || : ||| : | : | : |||
Db    110 TIENGFQIDGRARGGQSANDSYAKALNASSVQSALAVLREEPKDFVLQNHNIRSNL 169

      | : ||||| ||||| : ||||| : || : ||| : | : | : |||
Qy    61 DRIFDKTPEP 70
      | : ||| | |||
Db    170 ERIFAKAPEP 179

```

```

RESULT 2
US-08-838-151A-4
; Sequence 4, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; TITLE OF INVENTION: Genes
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838-151A

```

FILING DATE: 800  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mueller, Lisa V  
 REGISTRATION NUMBER: 38,978  
 REFERENCE/DOCKET NUMBER: SVS3801P0260  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-616-5400  
 TELEFAX: 312-616-5460  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 361 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PS-08-838-151A-4

Query Match 61.3%; Score 220; DB 4; Length 361;  
Best Local Similarity 58.6%; Pred. No. 2.3e-21;  
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps

QY 1 TLWGEFVDGRSARGCQTSDAAAEALNASSKEEALQIIRKIPPAALFPQHNLNSL 60  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 110 TI EWGFQIDGRSARGGQSANDSYAKALNASSVQSALAVLR EQPKDFVLQNHNIRSL 169  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 61 DRIFDKTPEP 70  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 170 ERIFA KAPEP 179

RESULT 3  
US-08-838-151A-6  
: Sequence 6, Application US/08838151A  
: Patent No. 6291743  
: GENERAL INFORMATION:  
: APPLICANT: Stout, John T  
: APPLICANT: Luu, Hang T

APPLICANT: Maxwell, Douglas  
APPLICANT: Ahlquist, Paul  
APPLICANT: Hanson, Steve  
TITLE OF INVENTION: Transgenic Plants Expressing Geninivirus  
TITLE OF INVENTION: Genes  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
STREET: Two Prudential Plaza, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/838,151A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mueller, Lisa V  
REGISTRATION NUMBER: 38,978  
REFERENCE/DOCKET NUMBER: SVS3801P0260  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5400  
TELEFAX: 312-616-5460  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-838-151A-6

```
Query Match          61.3%; Score 220; DB 4; Length 361;
Best Local Similarity 58.6%; Pred. No. 2.3e-21;
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps 0;
```

---

```
QY    1   TLVWGEFQVDGRSARGCGCOTSDNDAAEALNASSKEEALQIIREKIPAAALQFHNLNSL 60
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db    110 TIEMGFQIDGRSARGGQQSANDSYAKALNASSYQSVALVREOPKFVLQNHIRSNL 169

QY    61 DRIFDKTPEP 70
      :||| | |||
Db    170 ERIFAKAPEP 179
```

```

RESULT      4
US-08-838-151A-8
; Sequence 8, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
;   APPLICANT: Stout, John T
;   APPLICANT: Luu, Heng T
;   APPLICANT: Maxwell, Douglas
;   APPLICANT: Ahlquist, Paul
;   APPLICANT: Hanson, Steve
;   TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
;   TITLE OF INVENTION: Genes
;   NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dressler, Rocky, Milnamow & Katz
;   STREET: Two Prudential Plaza, Suite 4700
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: U.S.A.
;   ZIP: 60601
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk

```









US-08-838-151A-49

	Query Match	60.78;	Score 218;	DB 4;	Length 353;
	Best Local Similarity	60.00;	Pred. No. 4.le-21;		
	Matches 4;	Conservative 12;	Mismatches 16;	Indels	0; Gaps 0;
QY	1	TLVWGEFQDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALQFHNLNSL	60		
Db	110	TIEMGFQDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALQFHNLNSL	169		
QY	61	DRIFDKTPEP	70		
Db	170	ERIEVKVPEP	179		

```

RESULT 12
US-08-838-151A-52
; Sequence 52, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; TITLE OF INVENTION: Genes
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rockey, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A

```

[illegible]

RESULT 13  
 US-08-838-151A-55  
 ; Sequence 55, Application US/08838151A  
 ; Patent No. 6291743  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stout, John T  
 ; APPLICANT: Luu, Hang T  
 ; APPLICANT: Maxwell, Douglas  
 ; APPLICANT: Ahlquist, Paul  
 ; APPLICANT: Hanson, Steve  
 ; TITLE OF INVENTION: Transgenic Plants Expressing Genminivirus  
 ; TITLE OF INVENTION: Genes  
 ; NUMBER OF SEQUENCES: 63  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
 ; STREET: Two Prudential Plaza, Suite 4700  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: U.S.A.  
 ; ZIP: 60601  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/838,151A  
 ; FILING DATE:  
 ; CLASSIFICATION: 800  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mueller, Lisa V  
 ; REGISTRATION NUMBER: 38,978  
 ; REFERENCE/DOCKET NUMBER: SVS3801P0260  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-616-5400  
 ; TELEFAX: 312-616-5460  
 ; INFORMATION FOR SEQ ID NO: 55:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 353 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SVS-08-838-151A-55

```
Query Match          60.7% Score 218; DB 4; Length 353;
Best Local Similarity 60.0%; Pred. No. 4.le-21;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Qy      1 TLVNGEFQVGDSARGGCOTSDNAAAFALNASKEEALQIIREKIPAAALFQHNLNSL 60
       : ||| ||||| ||||| : ||| : ||| : ||| : ||| : ||| : |||
Db     110 TIENGQFQVGDSARGGQSANDSYAKALNADSTESALTILKEQPQDYVLQHHNIRSL 169

Qy      61 DRFDKTPPEP 70
       : ||| | |||
Db     170 ERIFVKVPPEP 179
```

```

RESULT 14
US-08-838-151A-24
; Sequence 24, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; TITLE OF INVENTION: Genes
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:

```

```

RESULT 15
US-08-838-151A-27
; Sequence 27, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dresser, Rockey, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A
; FILING DATE:
; CLASSIFICATION: 800

```

Search completed: January 3, 2002, 15:38:50  
Job time: 227 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:37:37 ; Search time 144.17 seconds  
(without alignments)  
35.965 Million cell updates/sec

Title: US-09-289-346A-6  
Perfect score: 357  
Sequence: 1 TLWGBFQVDRSARGCQT.....FAFALNSNLDRIKPTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	DB ID	Description
1	357	100.0	70 21 AAB18682	Mutant peptide der
2	340	95.2	70 21 AAB18677	Peptide fragment f
3	340	95.2	356 21 AAB18687	Amino acid sequenc
4	332	93.0	70 21 AAB18685	Mutant peptide der
5	329	92.2	70 21 AAB18688	Mutant peptide der
6	328	91.9	70 21 AAB18692	Mutant peptide der
7	326	91.3	70 21 AAB18684	Mutant peptide der
8	326	91.3	70 21 AAB18690	Mutant peptide der
9	325	91.0	70 21 AAB18678	Mutant peptide der
10	325	91.0	70 21 AAB18686	Mutant peptide der
11	324	90.8	70 21 AAB18689	Mutant peptide der

12	322	90.2	70 21 AAB18680	Mutant peptide der
13	322	90.2	70 21 AAB18691	Mutant peptide der
14	319	89.4	70 21 AAB18681	Mutant peptide der
15	318	89.1	70 21 AAB18683	Mutant peptide der
16	308	86.3	70 21 AAB18679	Mutant peptide der
17	210	58.8	359 17 AAR88870	Sardinian tomato y
18	210	58.8	359 17 AAR88871	Sardinian tomato y
19	210	58.8	359 17 AAR88872	Sardinian tomato y
20	206	57.7	353 18 AAW34338	Bean golden mosaic
21	206	57.7	353 18 AAW34332	Bean golden mosaic
22	206	57.7	353 18 AAW34333	Bean golden mosaic
23	206	57.7	353 18 AAW34334	Bean golden mosaic
24	206	57.7	353 18 AAW34335	Bean golden mosaic
25	204	57.1	353 8 AAP70407	ORF 4 gene product
26	204	57.1	361 18 AAW34336	Tomato mottle viru
27	204	57.1	361 18 AAW34324	Tomato mottle viru
28	204	57.1	361 18 AAW34325	Tomato mottle viru
29	204	57.1	361 18 AAW34326	Tomato mottle viru
30	197.5	55.3	361 8 AAP70562	Product of ORF 4 f
31	192	53.8	362 19 AAW56495	Tobacco leaf curl
32	191	53.5	357 18 AAW34329	Tomato yellow leaf
33	191	53.5	357 18 AAW34330	Tomato yellow leaf
34	191	53.5	357 18 AAW34331	Tomato yellow leaf
35	183	51.3	357 18 AAW34337	Tomato yellow leaf
36	67	18.8	447 21 AAY99659	Human GTPase assoc
37	67	18.8	447 21 AAY87089	Human secreted pro
38	67	18.8	447 21 AAY66744	Membrane-bound pro
39	67	18.8	447 21 AAY50947	Human adult aorta
40	67	18.8	447 22 AAE06066	Human gene 26 enco
41	67	18.8	447 22 AAB65267	Human PRO1125 (UNQ
42	67	18.8	456 21 AAY87190	Human secreted pro
43	67	18.8	456 22 AAE06167	Human gene 26 enco
44	64.5	18.1	1693 21 AAB48457	Human laminin 5 po
45	64.5	18.1	1693 21 AAB48459	Human laminin 5 po

## ALIGNMENTS

RESULT 1  
AAB18682  
ID AAB18682 standard; peptide; 70 AA.  
XX  
AC AAB18682;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; ALL; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 52 /note= "wild type residue replaced with Ala"  
FT Misc-difference 54 /note= "wild type residue replaced with Ala"  
FT Misc-difference 55 /note= "wild type residue replaced with Ala"  
FT  
XX WO200054573-A1.  
PN  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
PR  
XX 09-APR-1999; 99US-0289346.  
XX  
XX (UYN-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 XX WPI; 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 XX  
 XX Claim 53; Page 44-45; 73pp; English.  
 XX The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 XX Sequence 70 AA:  
 SQ  
 Query Match 100.0%; Score 357; DB 21; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-38;  
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSARGCQTSNDAAAEALNASSKEALQIIREKIPEKYLFAPALNSNL 60  
 Db 1 TLVWGEFQVDRSARGCQTSNDAAAEALNASSKEALQIIREKIPEKYLFAPALNSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70  
 RESULT 2  
 AAB18677  
 ID AAB18677 standard; peptide; 70 AA.  
 XX AAB18677:  
 AC  
 XX 22-JAN-2001 (first entry)  
 DT  
 XX Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
 DE  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 XX Tomato golden mosaic virus.  
 OS  
 XX WO200054573-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI; 2000-618851/59.  
 XX

PT Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 XX  
 XX Disclosure; Page 18; 73pp; English.  
 XX The present sequence is derived from a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX  
 XX Sequence 70 AA:  
 SQ  
 Query Match 95.2%; Score 340; DB 21; Length 70;  
 Best Local Similarity 95.7%; Pred. No. 8.9e-36;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSARGCQTSNDAAAEALNASSKEALQIIREKIPEKYLFAPALNSNL 60  
 Db 1 TLVWGEFQVDRSARGCQTSNDAAAEALNASSKEALQIIREKIPEKYLFQFHNSNL 60  
 QY 61 DRIFDKTPEP 70  
 Db 61 drifdktp 70  
 RESULT 3  
 AAB18687  
 ID AAB18687 standard; peptide; 356 AA.  
 XX AAB18687:  
 AC  
 XX 22-JAN-2001 (first entry)  
 DT  
 XX Amino acid sequence of a geminivirus replication protein of TGMV.  
 DE  
 XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 XX Tomato golden mosaic virus.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 354  
 FT /note="unspecified amino acid"  
 XX WO200054573-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000WO-US06759.  
 PF  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI; 2000-618851/59.  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT



PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 XX  
 PS Disclosure; Page 47-48; 73pp; English.  
 XX  
 CC The present sequence represents a geminivirus replication (Rep)  
 CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
 CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
 CC with other viral and host proteins. Mutants of the AL1 protein are used  
 CC to produce transgenic plants. The mutation in AL1 is present in a  
 CC ribosome binding region, and expression of mutant AL1 protein imparts  
 CC increased resistance to geminivirus infection in the plant. Mutant AL1  
 CC proteins are useful for producing plants having increased resistance or  
 CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
 CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
 CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
 CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
 CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
 CC beet curly top virus.  
 XX  
 SQ Sequence 356 AA;

Query Match 95.2%; Score 340; DB 21; Length 356;  
 Best Local Similarity 95.7%; Pred. No. 6.8e-35;  
 Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;..

OY 1 TLVWGFQVDGSRAGCGCQTSNDAAAEALNASSKEALQIIRKIPKYLFAFAALNSNL 60  
 |||||  
 Db 110 TLVWGFQVDGSRAGCGCQTSNDAAAEALNASSKEALQIIRKIPKYLFGFHlnsnl 169  
 |||||  
 OY 61 DRIFDKTPEP 70  
 |||||  
 Db 170 drifdktp 179

RESULT 4  
 AAB18685  
 ID AAB18685 standard; peptide; 70 AA.  
 XX  
 AC AAB18685;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 10 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 PN  
 XX PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI; 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT

PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 XX  
 PS Claim 53; Page 46; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;

Query Match 93.0%; Score 332; DB 21; Length 70;  
 Best Local Similarity 94.3%; Pred. No. 9.3e-35;  
 Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TLVWGFQVDGSRAGCGCQTSNDAAAEALNASSKEALQIIRKIPKYLFAFAALNSNL 60  
 |||||  
 Db 1 TLVWGFQVDGSRAGCGCQTSNDAAAEALNASSKEALQIIRKIPKYLFGFHlnsnl 60  
 |||||  
 OY 61 DRIFDKTPEP 70  
 |||||  
 Db 61 drifdktp 70

RESULT 5  
 AAB18688  
 ID AAB18688 standard; peptide; 70 AA.  
 XX  
 AC AAB18688;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX  
 KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX  
 OS Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 19 /note= "wild type residue replaced with Ala"  
 FT  
 XX WT Misc-difference 20 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 PN  
 XX PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06759.  
 XX  
 PR 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 PA (UYNC-) UNIV NORTH CAROLINA STATE.  
 XX  
 PI Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI; 2000-618851/59.  
 DR

XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
PS Disclosure; Page 48; 73pp; English.  
XX  
CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA;

Query Match 92.2%; Score 329; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 2.2e-34;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 TLWGFQVDGRSARGCQTSNDAAAEALNASSKEALQIIREKIPEKYLFAFAALNSNL 60  
|||||  
Db 1 tlwgfqvdgrsargcgcaasndaaaealnasskeaalqiirekipekylfqfhlnsl 60  
|||||

OY 61 DRIFDKTPEP 70  
|||||  
Db 61 drifdktp 70

RESULT 6  
AAB18692  
ID AAB18692 standard; peptide; 70 AA.  
XX  
AC AAB18692;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 66 /note= "wild type residue replaced with Ala"  
FT Misc-difference 69 /note= "wild type residue replaced with Ala"  
FT  
XX WO200054573-A1.  
XX  
XX PN  
XX PD 21-SEP-2000.  
XX  
XX PF 15-MAR-2000; 2000WO-US06759.  
XX  
XX PR 18-MAR-1999; 99US-0125004.  
XX PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.  
DR Transgenic plants with increased resistance to geminivirus infection  
XX comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
PS Disclosure; Page 50; 73pp; English.  
XX  
CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA;

Query Match 91.9%; Score 328; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 3e-34;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 TLWGFQVDGRSARGCQTSNDAAAEALNASSKEALQIIREKIPEKYLFAFAALNSNL 60  
|||||  
Db 1 tlwgfqvdgrsargcgctsndaaaealnasskeaalqiirekipekylfqfhlnsl 60  
|||||

OY 61 DRIFDKTPEP 70  
|||||  
Db 61 drifdktp 70

RESULT 7  
AAB18684  
ID AAB18684 standard; peptide; 70 AA.  
XX  
AC AAB18684;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 7 /note= "wild type residue replaced with Ala"  
FT Misc-difference 8 /note= "wild type residue replaced with Ala"  
FT  
XX WO200054573-A1.  
XX  
XX PN  
XX PD 21-SEP-2000.  
XX  
XX PF 15-MAR-2000; 2000WO-US06759.  
XX  
XX PR 18-MAR-1999; 99US-0125004.  
XX PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX PA

XX	
XX	Hanley-Bowdoin L, Orozco BM, Kong L;
XX	WPI; 2000-618851/59.
XX	
XX	Transgenic plants with increased resistance to geminivirus infection
PT	comprise a nucleic acid construct containing a nucleic acid sequence
PT	encoding a mutant All protein with a mutation in the Rb binding region
PT	-
XX	Claim 52; Page 45; 73pp; English.
PS	
XX	The present sequence represents a mutant peptide, derived from a
CC	geminivirus replication (Rep) protein, also known as All. All binds
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC	DNA, and interacts with other viral and host proteins. Mutants of the All
CC	protein are used to produce transgenic plants. The mutation in All is
CC	present in a ribosome binding region, and expression of mutant All
CC	protein imparts increased resistance to geminivirus infection in the
CC	plant. Mutant All proteins are useful for producing plants having
CC	increased resistance or reduced sensitivity to a geminivirus such as
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC	virus, cotton leaf curl virus or beet curly top virus.
XX	
XX	Sequence 70 AA:

Query Match 91.3%; Score 326; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 5.3e-34;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy	1	TLVWGFQVDGSRAGGCGTSDNAAAEALNASSKEEALQIIRKIKPEKYLFAFAALNSML	60
Db	1	tlvwgeaavdgsrsggqgsndaaaealnasskeeaqlirkeikekylfthlnsl	60
Qy	61	DRIFDKTPEP 70	
Db	61	drifdktpet 70	

RESULT	8	
AAAB18690		
ID	AAAB18690	standard; peptide; 70 AA.
XX	XX	
XX	AAAB18690;	
XX	XX	
DT	22-JAN-2001	(first entry)
XX	XX	
DE	DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX	XX	
KW	KW	Geminivirus; replication protein; Rep protein; ALL; transgenic plant;
KW	KW	ribosome binding region; resistance; geminivirus infection.
XX	XX	
OS	OS	Synthetic.
OS	OS	Tomato golden mosaic virus.

Key	Location/Qualifiers
Misc-difference 27	/note= "wild type residue replaced with Ala"
Misc-difference 30	/note= "wild type residue replaced with Ala"

XX	WO200054573-A1.
PN	
XX	
PD	21-SEP-2000.
XX	
PF	15-MAR-2000; 2000WO-US06759.
XX	
PR	18-MAR-1999; 99US-0125004.
PR	09-APR-1999; 99US-0289346.

XX	(UNYC-) UNIV NORTH CAROLINA STATE.
XX	XX PA
XX	XX PI
XX	XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX	XX WPI; 2000-618851/59.
XX	XX Transgenic plants with increased resistance to geminivirus infection
XX	XX comprise a nucleic acid construct containing a nucleic acid sequence
XX	XX encoding a mutant AL1 protein with a mutation in the Rb binding region
XX	XX -
XX	XX Disclosure; Page 49; 73pp; English.
XX	XX
XX	XX The present sequence represents a mutant peptide, derived from a
XX	XX geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX	XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX	XX DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX	XX protein are used to produce transgenic plants. The mutation in AL1 is
XX	XX present in a ribosome binding region, and expression of mutant AL1
XX	XX protein imparts increased resistance to geminivirus infection in the
XX	XX plant. Mutant AL1 proteins are useful for producing plants having
XX	XX increased resistance or reduced sensitivity to a geminivirus such as
XX	XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX	XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX	XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX	XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX	XX virus, cotton leaf curl virus or beet curly top virus.
XX	XX
XX	XX Sequence 70 AA: ,
XX	XX
XX	XX Query Match 91.3%; Score 326; DB 21: Length 70;
XX	XX Best Local Similarity 92.9%; Pred. No. 5.3e-34;
XX	XX Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX	XX
Qy	1 TLVWGEFQVQGRSARGGCQTSDNAAAEALNASSKEEALQITREKIPEKYLFAFAALNSNL 60
Db	1 tlvwgefvdgrsarggcqtsndaaaaaalaasskeaalqilrekipekylfqfhnlnsnl 60
Qy	61 DRIFDKTPEP 70
Db	61 drifdktpet 70
XX	XX
RESULT	9
AAB18678	
ID	AAB18678 standard; peptide; 70 AA.
XX	XX
AC	AAB18678;
XX	XX
DT	22-JAN-2001 (first entry)
XX	XX
DE	Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX	XX
KW	Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW	ribosome binding region; resistance; geminivirus infection.
XX	XX
OS	Synthetic.
OS	Tomato golden mosaic virus.
XX	XX
FH	Key Location/Qualifiers
FT	Misc-difference 12 /note= "wild type residue replaced with Ala"
FT	Misc-difference 13 /note= "wild type residue replaced with Ala"
FT	Misc-difference 15 /note= "wild type residue replaced with Ala"
XX	XX
PN	W0200054573-AL.
XX	XX
PD	21-SEP-2000.
XX	XX



```

FT XX /note= "wild type residue replaced with Ala"
PN WO200054573-A1.
XX XX
XX PD 21-SEP-2000.
XX XX
XX PF 15-MAR-2000; 2000WO-US06759.
XX XX
XX PR 18-MAR-1999; 99US-0125004.
XX PR 09-APR-1999; 99US-0289346.
XX XX
XX PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX DR WPI; 2000-618851/59.
XX XX
XX PT Transgenic plants with increased resistance to geminivirus infection
XX PT comprise a nucleic acid construct containing a nucleic acid sequence
XX PT encoding a mutant ALI protein with a mutation in the Rb binding region
XX PT
XX PS Disclosure; Page 48-49; 73pp; English.
XX CC
XX CC The present sequence represents a mutant peptide, derived from a
XX CC geminivirus replication (Rep) protein, also known as ALI. ALI binds
XX CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX CC DNA, and interacts with other viral and host proteins. Mutants of the ALI
XX CC protein are used to produce transgenic plants. The mutation in ALI is
XX CC present in a ribosome binding region, and expression of mutant ALI
XX CC protein imparts increased resistance to geminivirus infection in the
XX CC plant. Mutant ALI proteins are useful for producing plants having
XX CC increased resistance or reduced sensitivity to a geminivirus such as
XX CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX CC virus, cotton leaf curl virus or beet curly top virus.
XX CC
XX SQ Sequence 70 AA;

Query Match 90.8%; Score 324; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 9.6e-34;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGCGCQTSNDAAAEALNASSKEEALQIREKIPEKYLFAFAALNSNL 60
Db 1 tlwgefvdgrsarggcqtsndaaaealnasskeaalqirekipekylfqfhnlnsl 60

QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70

RESULT 12
AAB18680
XX AAB18680 standard; peptide; 70 AA.
XX AAB18680;
XX XX
XX DT 22-JAN-2001 (first entry)
XX DE Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.
XX KW Geminivirus; replication protein; Rep protein; ALI; transgenic plant;
XX KW ribosome binding region; resistance; geminivirus infection.
XX OS Synthetic.
XX OS Tomato golden mosaic virus.
XX FH Key Location/Qualifiers
FT Misc-difference 42

```

```

FT Misc-difference 43 /note= "wild type residue replaced with Ala"
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 44 /note= "wild type residue replaced with Ala"
XX WO200054573-A1.
XX 21-SEP-2000.
XX 15-MAR-2000; 2000WO-US06759.
XX 18-MAR-1999; 99US-0125004.
XX 09-APR-1999; 99US-0289346.
XX (UYNC-) UNIV NORTH CAROLINA STATE.
XX Hanley-Bowdoin L, Orozco BM, Kong L;
XX WPI; 2000-618851/59.
XX Transgenic plants with increased resistance to geminivirus infection
XX comprise a nucleic acid construct containing a nucleic acid sequence
XX encoding a mutant ALI protein with a mutation in the Rb binding region
XX
XX Claim 52; Page 43-44; 73pp; English.
XX The present sequence represents a mutant peptide, derived from a
XX geminivirus replication (Rep) protein, also known as ALI. ALI binds
XX double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX DNA, and interacts with other viral and host proteins. Mutants of the ALI
XX protein are used to produce transgenic plants. The mutation in ALI is
XX present in a ribosome binding region, and expression of mutant ALI
XX protein imparts increased resistance to geminivirus infection in the
XX plant. Mutant ALI proteins are useful for producing plants having
XX increased resistance or reduced sensitivity to a geminivirus such as
XX tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX virus, cotton leaf curl virus or beet curly top virus.
XX
XX SQ Sequence 70 AA;

Query Match 90.2%; Score 322; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 1.7e-33;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSGCGCQTSNDAAAEALNASSKEEALQIREKIPEKYLFAFAALNSNL 60
Db 1 tlwgefvdgrsarggcqtsndaaaealnasskeaalqialaaip 60

QY 61 DRIFDKTPEP 70
Db 61 drifdktp 70

RESULT 13
AAB18691
XX ID AAB18691 standard; peptide; 70 AA.
XX AAB18691;
XX XX
XX DT 22-JAN-2001 (first entry)
XX DE Mutant peptide derived from amino acids 110-179 of Rep (ALI) protein.
XX KW Geminivirus; replication protein; Rep protein; ALI; transgenic plant;
XX KW ribosome binding region; resistance; geminivirus infection.
XX OS Synthetic.

```



22-JAN-2001 (first entry)

Search completed: January 3, 2002, 15:37:37





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:57:16 ; Search time 43.68 seconds  
(without alignments)  
58.758 Million cell updates/sec

Title: US-09-289-346A-6

Perfect score: 357

Sequence: 1 TLWGFQVQDGRSARGCQT.....FAFAALNSLDRIFDKTPEP 70

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	340	95.2	352	1 VAL1_TGMV	P03567 tomato gold
2	240	67.2	361	1 VAL1_PYMV	P27258 potato yell
3	227	63.6	358	1 VAL1_CLVK	P14982 cassava lat
4	227	63.6	358	1 VAL1_CLVN	P14972 cassava lat
5	220	61.6	362	1 VAL1_TYLC	P36279 tomato yell
6	217	60.8	359	1 VAL1_TYLC	P38609 tomato yell
7	210	58.8	359	1 VAL1_TYLC	P27260 tomato yell
8	209	58.5	349	1 VAL1_PHUV	P06923 pepper huas
9	204	57.1	353	1 VAL1_BGMV	P05175 bean golden
10	203	56.9	355	1 VAL1_ABMVW	P21947 abutilon mo
11	200	56.0	358	1 VAL1_BCTV	P14991 beet curly
12	200	56.0	357	1 VAL1_TMOV	P06657 tomato molt
13	191	53.5	357	1 VAL1_TYLC	P27259 tomato yell
14	104	29.1	347	1 VAL1_SLCV	P29048 squash leaf
15	67	18.8	447	1 TBL2_HUMAN	Q9V4P3 homo sapien
16	64.5	18.1	1713	1 LMA3_HUMAN	Q16787 homo sapien
17	63.5	17.8	1610	1 CCAD_MESAU	Q99244 mesocricetu
18	63.5	17.8	2161	1 CCAD_HUMAN	Q01668 homo sapien
19	63.5	17.8	2203	1 CCAD_RAT	P27732 rattus norv
20	60	16.8	269	1 ORN_YEAST	P54964 saccharomyc
21	60	16.8	295	1 VAL1_TYDA	P31617 tobacco yell
22	60	16.8	355	1 CRTB_RHOSH	P54905 rhodobacte
23	59.5	16.7	299	1 Y175_HELPJ	Q9ZM97 helicobacte
24	59.5	16.7	470	1 RISA_RHIME	Q9Z3R2 rhizobium m
25	59.5	16.7	4385	1 YP73_CAEEL	Q9ZED6 caenorhabdi
26	58.5	16.4	706	1 Y006_RICPR	Q9ZED6 rickettsia
27	58.5	16.4	1232	1 Y005_CAEEL	P34643 caenorhabdi
28	58	16.2	207	1 IL6_MARMO	O35736 marmota mon
29	58	16.2	447	1 CDSA_DROME	P56079 d phosphati
30	57.5	16.1	367	1 LHX4_MOUSE	P53776 mus musculu
31	57.5	16.1	844	1 SECA_STACA	P47994 staphylococ
32	57.5	16.1	1852	1 CCAS_CYPCA	P22316 cyprinus ca
33	57	16.0	1483	1 CYP1_YEAST	P12351 saccharomyc

34 57 16.0 1608 1 HLYA\_SERMA P15320 serratia ma  
35 56.5 15.8 511 1 HUTH\_VIBCH Q9Ksq4 vibrio chol  
36 56.5 15.8 2190 1 CCAD\_CHICK O73700 gallus gall  
37 56 15.7 258 1 YL48\_MYCTU O06228 mycobacteri  
38 56 15.7 408 1 CINA\_THEMA Q985X1 thermotoga  
39 56 15.7 492 1 MOT3\_MOUSE O35308 mus musculu  
40 56 15.7 492 1 MOT3\_RAT O70461 rattus norv  
41 56 15.7 703 1 GYS2\_RAT P17625 rattus norv  
42 56 15.7 767 1 TOP1\_MOUSE O04750 mus musculu  
43 56 15.7 1033 1 YDK9\_SCHPO P87115 schizosacch  
44 56 15.7 1051 1 CARB\_SULSO Q59969 sulfobobus  
45 56 15.7 1663 1 HAPD\_HUMAN O60229 homo saplen

#### ALIGNMENTS

RESULT 1  
VAL1\_TGMV  
ID VAL1\_TGMV STANDARD; PRT; 352 AA.  
AC P03567;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL1 PROTEIN.  
GN AC1.  
OS Tomato golden mosaic virus (TGMV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10831;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;  
RT "Complete" nucleotide sequence of the infectious cloned DNA components  
RT of tomato golden mosaic virus: potential coding regions and regulatory  
RT sequences.  
RL EMBO J. 3:2197-2205(1984).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; K02029; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A04170; QCCVLI.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATALL.  
DR PRINTS; PR00228; GEMCOATCLVLL.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
KW ATP-binding. 223 230 ATP (POTENTIAL).  
FT NP\_BIND 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;  
SQ SEQUENCE

Query Match 95.2%; Score 340; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 1.5e+30;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TLWGFQVQDGRSARGCQTSDNDAAAFALNSKKEALQIIRKIPKYLFAFAALNSNL 60

DB 111 TLWGFQVQDGRSARGCQTSDNDAAAFALNSKKEALQIIRKIPKYLFAFAALNSNL 170

OY 61 DRIFDKTPEP 70

DB 171 DRIFDKTPEP 180

RESULT 2

VAL1\_PYMV

ID VALL\_PYMV STANDARD; PRT; 361 AA.  
AC P27258;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE ALL PROTEIN.  
OS Potato yellow mosaic virus (isolate Venezuela).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10828;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91311403; PubMed=1856690;  
RA Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;  
RT "The nucleotide sequence of the infectious cloned DNA components of potato yellow mosaic virus.";  
RL J. Gen. Virol. 72:1515-1520(1991).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: D00940; BAA00782.1;  
DR PIR: J00364; QOCVPT.  
DR InterPro: IPR001191; Geminini\_AL1.  
DR Pfam: PF00799; Geminini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR PRINTS: PR00228; GEMCOATCLVL1.  
DR ProDom: PD000736; Geminini\_AL1; 1.  
KW ATP-binding. 222 229 ATP (POTENTIAL).  
FT NP\_BIND 361 AA; 40850 MW; 5627A33BF1264383 CRC64;  
SQ SEQUENCE 361 AA; 40850 MW; 5627A33BF1264383 CRC64;

Query Match 67.2%; Score 240; DB 1; Length 361;  
Best Local Similarity 66.7%; Pred. No. 26-19; Indels 0; Gaps 0;  
Matches 46; Conservative 11; Mismatches 12;  
QY 1 TLVWGFQVDSARGGCGTSDAAAEALNASSKEALQIIREKIPKYLFAFAALNSNL 60  
Db 110 TLVWGLFQIDGRSARGGQTVNDAAAEALNSGTKEAAMKIIEKLPKELFQYHNLSNL 169  
QY 61 DRIFDKTPE 69  
Db 170 DRIFMKAPE 178

RESULT 3  
VALL\_CLVK  
ID VALL\_CLVK STANDARD; PRT; 358 AA.  
AC P14982;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE ALL PROTEIN (40.4 KDA PROTEIN).  
GN AC1.  
OS Cassava latent virus (strain West Kenyan 844).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stanley J., Gay M.R.;  
RT "Nucleotide sequence of cassava latent virus DNA.";  
RL Nature 301:260-262(1983).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: X17095; CAA34953.1;  
DR PIR: S07594; S07594.  
DR InterPro: IPR001191; Geminini\_AL1.  
DR Pfam: PF00799; Geminini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR PRINTS: PR00228; GEMCOATCLVL1.  
DR ProDom: PD000736; Geminini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 220 227 ATP (POTENTIAL).  
SQ SEQUENCE 358 AA; 40435 MW; 1DB16BB0CB2D5E2C CRC64;

Query Match 63.6%; Score 227; DB 1; Length 358;  
Best Local Similarity 60.0%; Pred. No. 5.6e-18;  
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: J02057; NOT ANNOTATED CDS.  
DR InterPro: IPR001191; Geminini\_AL1.  
DR Pfam: PF00799; Geminini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR PRINTS: PR00228; GEMCOATCLVL1.  
DR ProDom: PD000736; Geminini\_AL1; 1.  
KW ATP-binding. 220 227 ATP (POTENTIAL).  
FT NP\_BIND 358 AA; 40346 MW; ED173E753EE92D69 CRC64;  
SQ SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 63.6%; Score 227; DB 1; Length 358;  
Best Local Similarity 60.0%; Pred. No. 5.6e-18;  
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TLVWGFQVDSARGGCGTSDAAAEALNASSKEALQIIREKIPKYLFAFAALNSNL 60  
Db 109 TLVWGLFQIDGRSARGGQTVNDAAAEALNSGTKEAAMKIIEKLPKELFQYHNLSNL 168  
QY 61 DRIFDKTPE 70  
Db 169 DRIFQEPAP 178

RESULT 4  
VALL\_CLVN  
ID VALL\_CLVN STANDARD; PRT; 358 AA.  
AC P14972;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE ALL PROTEIN (40.4 KDA PROTEIN).  
GN AC1.  
OS Cassava latent virus (strain Nigerian).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10819;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90174930; PubMed=2308831;  
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.;  
RT "Nucleotide sequence of the infectious cloned DNA components of African cassava mosaic virus (Nigerian strain).";  
RL Nucleic Acids Res. 18:197-198(1990).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: X17095; CAA34953.1;  
DR PIR: S07594; S07594.  
DR InterPro: IPR001191; Geminini\_AL1.  
DR Pfam: PF00799; Geminini\_AL1; 1.  
DR PRINTS: PR00227; GEMCOATALL.  
DR PRINTS: PR00228; GEMCOATCLVL1.  
DR ProDom: PD000736; Geminini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 220 227 ATP (POTENTIAL).  
SQ SEQUENCE 358 AA; 40435 MW; 1DB16BB0CB2D5E2C CRC64;

Query Match 63.6%; Score 227; DB 1; Length 358;  
Best Local Similarity 60.0%; Pred. No. 5.6e-18;  
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

RA	Reis E., Rodrigo E., Acosta G., Wolfones E.,	DR	PLAM: PR00799; Gemini_Atl: 1.
RT	"High similarity among the tomato yellow leaf curl virus isolates	DR	PRINTS: PR00227; GEMCOATL1.
RT	from the west Mediterranean basin: the nucleotide sequence of an	DR	PRINTS: PR00228; GEMCOATCLV1.
RT	infectious clone from Spain.";	DR	PRIDom: PD000736; Gemini_Atl: 1.

KW	ATP-binding.	220	227	ATP (POTENTIAL).	
FT	NP_BIND	359	AA:	40733 MW; 9717B4A07C93EFA7 CRC64;	
SQ	SEQUENCE	359	AA:	40733 MW; 9717B4A07C93EFA7 CRC64;	
Query Match					
Best Local Similarity					
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;					
QY	2	LVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD	61		
DB	111	LENGTQIDGRSARGCGQTANDAYAKAINAGSKSALDVKEIAPDRDVLVHFNHNSNLD	170		
QY	62	RIFDKTPEP 70			
DB	171	KVFQVPPAP 179			
RESULT 8					
ID	VALI_PHVU	STANDARD;	PRT;	349	AA.
AC	Q06923;				
DT	01-JUN-1994 (Rel. 29, Created)				
DT	01-JUN-1994 (Rel. 29, Last sequence update)				
DT	01-JUN-1994 (Rel. 29, Last annotation update)				
DE	ALL PROTEIN.				
GN	ALL.				
OS	Pepper huasteco virus (PHV).				
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.				
OX	NCBI_TaxID=28349;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Torres-Pacheco I., Garzon-Riznado J.A., Herrera-Estrella L., Rivera-Bustamante R.F.;				
RA	"Complete nucleotide sequence of pepper huasteco virus: analysis and comparison with bipartite geminiviruses.";				
RT	J. Gen. Virol. 74:2225-2231(1993).				
RL	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL: X70418; CAA49856.1; -				
DR	PIR: S31875; S31875.				
DR	PIR: J02300; J02300.				
DR	InterPro: IPR001191; Gemini_ALL.				
DR	Pfam: PF00799; Gemini_ALL; 1.				
DR	PRINTS: PR00227; GEMCOATALL.				
DR	PRINTS: PR00228; GEMCOATCLVL1.				
DR	ProDom: PD000736; Gemini_ALL; 1.				
FT	ATP-binding. 221 228				
FT	NP_BIND				
SQ	SEQUENCE 349 AA; 39722 MW; D5F4E76CD56370F4 CRC64;				
Query Match					
Best Local Similarity					
Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;					
QY	1	TLWVGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD	60		
DB	110	TVWGEFQVDRSGRGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD	169		
QY	61	DRIFDKTPEP 70			
DB	170	NRIFQVPPAP 179			
RESULT 9					
ID	VALI_BGMV	STANDARD;	PRT;	353	AA.
AC	P05175;				
DT	13-AUG-1987 (Rel. 05, Created)				
DT	13-AUG-1987 (Rel. 05, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	ALL PROTEIN (40.2 KDA PROTEIN).				
GN	AC1.				
OS	Bean golden mosaic virus.				
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.				
OX	NCBI_TaxID=10839;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Huwarth A.J., Caton J., Bossert M., Goodman R.M.;				
RA	"Nucleotide sequence of bean golden mosaic virus and a model for gene regulation in geminiviruses.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).				
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL: M10070; AAA46318.1; -				
DR	InterPro: IPR001191; Gemini_ALL.				
DR	Pfam: PF00799; Gemini_ALL; 1.				
DR	PRINTS: PR00227; GEMCOATALL.				
DR	PRINTS: PR00228; GEMCOATCLVL1.				
DR	ProDom: PD000736; Gemini_ALL; 1.				
KW	ATP-binding. 222 229				
FT	NP_BIND				

```
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR PIR; A36214; QQCW1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP-BIND 221 228 ATP (POTENTIAL).
SQ SEQUENCE 355 AA; 40257 MW; 16A2CA8A63251E95 CRC64;

Query Match 56.9%; Score 203; DB 1; Length 355;
Best Local Similarity 55.7%; Pred. No. 2.6e-15;
Matches 39; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
DB 110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVOSALNILEEQPKDYFLQNHNRSL 169

QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179

RESULT 11
VAL1_BCTV STANDARD; PRT; 358 AA.
AC P14991;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE ALL PROTEIN (40.8 KDA PROTEIN).
OS Beet curly top virus (BCTV).
OC Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX NCBI_TaxID=10840;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT "The nucleotide sequence of an infectious clone of the geminivirus
RL EMBO J. 5:1761-1767(1986).
CC 1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP-BIND 222 229 ATP (POTENTIAL).
SQ SEQUENCE 358 AA; 40889 MW; 39A45FE3C0B9C333 CRC64;

Query Match 56.0%; Score 200; DB 1; Length 358;
Matches 39; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
DB 110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVOSALNILEEQPKDYFLQNHNRSL 169

QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179

RESULT 12
VAL1_TMOV STANDARD; PRT; 361 AA.
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN ALL.
OS Tomato mottle virus (isolate Florida) (TMov).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=36449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107858; PubMed=1469361;
RA Abouzid A.M., Polston J.E., Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RL J. Gen. Virol. 73:3225-3229(1992).
CC 1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14460; AAC32414.1; -
DR PIR; JQ1870; JQ1870.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP-BIND 222 229 ATP (BY SIMILARITY).
SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEEAC6950 CRC64;

Query Match 56.0%; Score 200; DB 1; Length 361;
Best Local Similarity 52.9%; Pred. No. 5.7e-15;
Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
DB 110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALILKEEQPKDYFLQHHNLNA 169

QY 61 DRIFDKTPEP 70
DB 170 QKIFORPPDP 179

RESULT 13
VAL1_TYLCV STANDARD; PRT; 357 AA.
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
```

```
-----
Best Local Similarity 52.9%; Pred. No. 5.6e-15;
Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
DB 110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALILKEEQPKDYFLQHHNLNA 169

QY 61 DRIFDKTPEP 70
DB 170 QKIFORPPDP 179

RESULT 12
VAL1_TMOV STANDARD; PRT; 361 AA.
AC Q06657;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE ALL PROTEIN.
GN ALL.
OS Tomato mottle virus (isolate Florida) (TMov).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=36449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107858; PubMed=1469361;
RA Abouzid A.M., Polston J.E., Hiebert E.;
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus
RL J. Gen. Virol. 73:3225-3229(1992).
CC 1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14460; AAC32414.1; -
DR PIR; JQ1870; JQ1870.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATALL.
DR PRINTS; PR00228; GEMCOATCLVL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP-BIND 222 229 ATP (BY SIMILARITY).
SQ SEQUENCE 361 AA; 40516 MW; 8138B65CEEAC6950 CRC64;

Query Match 56.0%; Score 200; DB 1; Length 361;
Best Local Similarity 52.9%; Pred. No. 5.7e-15;
Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDSRGSGGCGTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
DB 110 TIEWGEFQIDGRSARGGQQTANDSYAKALNAGVSQALAVLREEQPKDYFLQHHNRSL 169

QY 61 DRIFDKTPEP 70
DB 170 ERIFAKAPEP 179

RESULT 13
VAL1_TYLCV STANDARD; PRT; 357 AA.
AC P27259;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
```

DE ALL PROTEIN (C1 PROTEIN).  
GN C1.  
OS Tomato yellow leaf curl virus (TYLCV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10832;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92024070; PubMed=1926771;  
RA Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;  
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus  
with a single genomic component.";  
RL Virology 185:151-161(1991).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X15656; CA333688.1; .  
DR PIR; D40779; QOCVCL.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR PRINTS; PR00228; GEMCOATCLVL1.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 219 226 ATP (POTENTIAL).  
SQ SEQUENCE 357 AA; 40678 MW; 939AB6BE1AB3B2A7 CRC64;  
  
Query Match 53.5%; Score 191; DB 1; Length 357;  
Best Local Similarity 62.3%; Pred. No. 5.6e-14;  
Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;  
  
QY 4 WGEFQVGRSARGCQTSNDAAAEALNASKKEALQIIREKIPEKYLFAFAALNSNLDRI 63  
DB 111 FGVSGIDGRSARGGQGSANDAYAEALNSGSKSEALILKEKAPKDYILOFHNLSNLDRI 170  
QY 64 F 64  
DB 171 F 171  
  
RESULT 14  
ID VAL1\_SLCV STANDARD; PRT; 347 AA.  
AC P29048;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-DEC-1992 (Rel. 24, Last annotation update)  
DE ALL PROTEIN.  
OS Squash leaf curl virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10829;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91082449; PubMed=1984668;  
RA Lazarowitz S.G., Lazdins I.B.;  
RT "Infectivity and complete nucleotide sequence of the cloned genomic  
components of a bipartite squash leaf curl geminivirus with a broad  
host range phenotype.";  
RL Virology 180:58-69(1991).  
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES ALL PROTEIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; M38183; AAC32410.1; ALT\_INIT.  
DR PIR; C36785; QOCVSL.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR PRINTS; PR00228; GEMCOATCLVL1.  
DR ProDom; PD000736; Gemini\_AL1; 1.  
KW ATP-binding.  
FT NP\_BIND 218 225 ATP (POTENTIAL).  
SQ SEQUENCE 347 AA; 39110 MW; AFDAEBDE122110E CRC64;  
  
Query Match 29.1%; Score 104; DB 1; Length 347;  
Best Local Similarity 34.8%; Pred. No. 0.00025;  
Matches 23; Conservative 14; Mismatches 25; Indels 4; Gaps 1;  
  
QY 5 GFEQVGRSARGCQTSNDAAAEALNASKKEALQIIREKIPEKYLFAFAALNSNLDRI 64  
DB 116 GQYKVSQ-----GSKSNKDDVYHNAVAGSAGEALDIKAGDPKTFIVYHNLNLANVERLF 171  
QY 65 DKTPEP 70  
DB 172 QKPPEP 177  
  
RESULT 15  
ID TBL2\_HUMAN STANDARD; PRT; 447 AA.  
AC Q9Y4P3; Q9UQE2;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TRANSDUCIN BETA-LIKE 2 PROTEIN (WS BETA-TRANSDUCIN REPEATS PROTEIN)  
DE (WS-BETATRP).  
GN TBL2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20044626; PubMed=10575226;  
RA Perez Jurado L.A., Wang Y.-K., Francke U., Cruces J.;  
RT "TBL2, a novel transducin family member in the WBS deletion:  
characterization of the complete sequence, genomic structure,  
transcriptional variants and the mouse ortholog.";  
RL Cytogenet. Cell Genet. 86:277-284(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=21154917; PubMed=11230166;  
RA Wiemann S., Weil B., Wellenreuther R., Cassenhuber J., Glassl S.,  
RA Ansorge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
RA Newes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
RA Wambutt R., Korn B., Klein M., Poustka A.;  
RT "Towards a catalog of human genes and proteins: sequencing and  
analysis of 500 novel complete protein coding human cDNAs.";  
RL Genome Res. 11:422-435(2001).  
RN [3]  
RP SEQUENCE OF 14-447 FROM N.A.  
RX MEDLINE=99075645; PubMed=9860302;  
RA Meng X., Lu X., Li Z., Green E.D., Massa H., Trask B.J., Morris C.A.,  
RA Keating M.T.;  
RT "Complete physical map of the common deletion region in Williams  
syndrome and identification and characterization of three novel  
genes.";  
RL Hum. Genet. 103:590-599(1998).  
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).  
CC

Search completed: January 3, 2002, 15:57:17  
Job time: 1103 sec

---

**THIS PAGE BLANK (USPTO)**





```

009727
ID 009727; PRELIMINARY; PRT; 226 AA.
AC 009727;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR InterPro; IPR001191; Gemin_Ali.
DR Pfam; PF00799; GEMCOATL1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_Ali; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 76.2%; Score 272; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 7.8e-23;
Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLVWGFQVDSRGSGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
   1: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111 TVWGFQVDSRGSGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFOYHNLSSNL 170
   1: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 DRIFDKTPEP 70
   ||||| |||||
Db 171 DRIFAKAPEP 180

RESULT 3
Q9WHF6 PRELIMINARY; PRT; 226 AA.
AC Q9WHF6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN [1]
RC STRAIN=HN96-H5KW;
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M., Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131071; AAD33471.1; -.
DR InterPro; IPR001191; Gemin_Ali.
DR Pfam; PF00799; Gemin_Ali; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_Ali; 1.
FT NON_TER 226
SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 76.2%; Score 272; DB 12; Length 226;
Best Local Similarity 72.9%; Pred. No. 7.8e-23;
Matches 51; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 TLVWGFQVDSRGSGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
   1: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 111 TIWGFQVDSRGSGCQTSDNDAAEALNASSKEEAMRIKEKLPKLFQYHNLSSNL 170

QY 61 DRIFDKTPEP 70
   ||||| |||||
Db 171 DRIFAKAPEP 180

RESULT 4
Q67574 PRELIMINARY; PRT; 361 AA.
AC Q67574;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE REPLICATIVE PROTEIN.
GN AL1.
OS Bean golden mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10839;
RN [1]
RC SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G., Maxwell D.P., Russell D.R.;
RL Plant Dis. 75:336-342(1991).
RN [2]
RN [3]
RP SEQUENCE FROM N.A.
RA Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M88686; AAA46312.1; -.
DR InterPro; IPR001191; Gemin_Ali.
DR Pfam; PF00799; Gemin_Ali; 1.
DR PRINTS; PR00227; GEMCOATL1.
DR ProDom; PD000736; Gemin_Ali; 1.
SQ SEQUENCE 361 AA; 41041 MW; 0094C7ACAF06B788 CRC64;

Query Match 75.6%; Score 270; DB 12; Length 361;
Best Local Similarity 76.1%; Pred. No. 2.3e-22;
Matches 51; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 4 WGFQVDSRGSGCQTSDNDAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
   1: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 WGFQVDSRGSGCQTSDNDAAEALNASSKEEAMRIKEKLPKLFQYHNLSSNL 172

QY 64 FDKTPEP 70
   ||||| |||||
Db 173 FTRAPDP 179

RESULT 5
Q9QDB1 PRELIMINARY; PRT; 225 AA.
AC Q9QDB1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS cowpea golden mosaic geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=69263;
RN [1]
RC SEQUENCE FROM N.A.
RA Faria J.C.;
RT "Partial nucleotide sequence of cowpea golden mosaic geminivirus from Brazil.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AF188708; AAF06318.1; -.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR PRODOM; PD000736; Gemini\_AL1; 1.  
FT NON\_TER 225  
SQ SEQUENCE 225 AA; 25766 MW; 1089CB6B8D15B5D CRC64;

Query Match 75.1%; Score 268; DB 12; Length 225;  
Best Local Similarity 76.1%; Pred. No. 2.2e-22;  
Matches 51; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 WGEFQVQDGRSARGCGCOTSDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNLDRI 63  
|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||  
Db 113 WGHFQIDGRSARGGQQTINDAAEALNASSKEEAMQIIKEKLPKFLFYHNLSSNLDRI 172

QY 64 FDKTPEP 70  
||| | |||  
Db 173 FRKPPPEP 179

RESULT 6  
Q9ELT8 PRELIMINARY; PRT; 314 AA.  
AC Q9ELT8;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE REPLICATION ASSOCIATION PROTEIN.  
GN AC1.

OS sweet potato leaf curl virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=100755;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
RT "Detection of a geminivirus infecting sweet potato in the United  
RT States";  
RL Plant Dis. 82:1253-1257(1998).  
RN [2]

RP SEQUENCE FROM N.A.  
RA Lotrakul P., Valverde R.A.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF288227; AAG01006.1; -.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR PRODOM; PD000736; Gemini\_AL1; 1.  
SQ SEQUENCE 314 AA; 35153 MW; 6B6220613046943F CRC64;

Query Match 75.1%; Score 268; DB 12; Length 314;  
Best Local Similarity 79.4%; Pred. No. 3.2e-22;  
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVQDGRSARGCGCOTSDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
|:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||  
Db 110 TITWGEFQVQDGRSARGGQQTANDAAEALNAGSKEAALQIIREKLPKFLFYHNLVSNL 169

QY 61 DRIFDKTP 68  
|||| |  
Db 170 DRIFSPPP 177

RESULT 7  
Q9QS55 PRELIMINARY; PRT; 364 AA.  
AC Q9QS55;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE REPLICATION INITIATION PROTEIN AC1.

GN AC1.  
OS sweet potato leaf curl virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=100755;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;  
RT "Detection of a geminivirus infecting sweet potato in the United  
RT States";  
RL Plant Dis. 82:1253-1257(1998).  
DR EMBL; AF104036; AAD47173.1; -.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR PRODOM; PD000736; Gemini\_AL1; 1.  
SQ SEQUENCE 364 AA; 40680 MW; 5F79752431A09D6E CRC64;

Query Match 75.1%; Score 268; DB 12; Length 364;  
Best Local Similarity 79.4%; Pred. No. 3.8e-22;  
Matches 54; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 TLVWGEFQVQDGRSARGCGCOTSDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
|:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||  
Db 110 TITWGEFQVQDGRSARGGQQTANDAAEALNAGSKEAALQIIREKLPKFLFYHNLVSNL 169

QY 61 DRIFDKTP 68  
|||| |  
Db 170 DRIFSPPP 177

RESULT 8  
Q98693 PRELIMINARY; PRT; 185 AA.  
AC Q98693;  
DT 01-FEB-1997 (Tremblrel. 02, Created)  
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE REP PROTEIN (FRAGMENT).  
GN AC1.

OS sida golden mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=51034;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAMAICA;  
RA Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;  
RL Plant Dis. 81:1251-1258(1997).  
DR EMBL; U67926; AAB97865.1; -.  
DR InterPro; IPR001191; Gemini\_AL1.  
DR Pfam; PF00799; Gemini\_AL1; 1.  
DR PRINTS; PR00227; GEMCOATL1.  
DR PRODOM; PD000736; Gemini\_AL1; 1.  
FT NON\_TER 1  
FT NON\_TER 185  
SQ SEQUENCE 185 AA; 20975 MW; 3913850A025A5EE1 CRC64;

Query Match 74.2%; Score 265; DB 12; Length 185;  
Best Local Similarity 70.0%; Pred. No. 3.8e-22;  
Matches 49; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

QY 1 TLVWGEFQVQDGRSARGCGCOTSDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
|:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||  
Db 89 TIWGEFQIDGRSARGGQQTANDAAAEALNAGSGTKEDALKIIREKLPKFLFYHNLSSNI 148

QY 61 DRIFDKTPEP 70  
|||| | |||  
Db 149 DRIFSKPPEP 158

RESULT 9  
P88975

ID P88975 PRELIMINARY; PRT: 149 AA.  
 AC P88975;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 GN AC1.  
 OS Macroptilium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA;  
 RA Roye M.E.;  
 RL Thesis (1996). Biochemistry. University of the West Indies, Jamaica.  
 DR EMBL: U75278; AAB36919.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 1  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 16785 MW; E4CF5BED4C9CD508 CRC64;

Query Match 72.0%; Score 257; DB 12; Length 149;  
 Best Local Similarity 68.6%; Pred. No. 2.3e-21;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDRSARGGCGTSDNAAAEALNASKEPALQIIREKIPEKYLFAFAALNSNL 60  
 Db 52 TIWGVFQIDGRSARGGCGTSDNAAAEALNSGTKEAAMRIVKEKLPKELFQYHNLSSNL 111  
 QY 61 DRIFDKTPEP 70  
 Db 112 DRIFMKDPEP 121

RESULT 10  
 O9YLA4 PRELIMINARY; PRT: 233 AA.  
 ID O9YLA4;  
 AC O9YLA4;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION ASSOCIATED PROTEIN (FRAGMENT).  
 GN REP.  
 OS Macroptilium golden mosaic geminivirus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=51676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E.;  
 RT "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses from Jamaica.";  
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAMAICA STRAIN 1;  
 RA Roye M.E., McLaughlin W.A., Maxwell D.P.;  
 RT "Molecular characterization of two distinct geminiviruses infecting M. lathyroides from Jamaica.";  
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF098940; AADI7850.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 233 233  
 SQ SEQUENCE 233 AA; 26355 MW; AA490AF4D2166A02 CRC64;

Query Match 72.0%; Score 257; DB 12; Length 233;  
 Best Local Similarity 68.6%; Pred. No. 3.9e-21;  
 Matches 48; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDRSARGGCGTSDNAAAEALNASKEPALQIIREKIPEKYLFAFAALNSNL 60  
 Db 110 TIWGVFQIDGRSARGGCGTSDNAAAEALNSGTKEAAMRIVKEKLPKELFQYHNLSSNL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 DRIFMKDPEP 179  
 RESULT 11  
 O39180 PRELIMINARY; PRT: 234 AA.  
 ID O39180;  
 AC O39180;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).  
 OS potato yellow mosaic virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10827;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TOMATO STRAIN;  
 RA Guzman P., Arredondo C.R., Emmatty D., Portillo R.J.; Gilbertson R.L.;  
 RL Plant Dis. 81:312-312(1997).  
 DR EMBL: AF026553; AAB82605.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.  
 DR PRINTS: PR00227; GEMCOATALL.  
 DR ProDom: PD000736; Gemini\_AL1; 1.  
 FT NON\_TER 234 234  
 SQ SEQUENCE 234 AA; 26486 MW; 9ED8F0697105CD19 CRC64;

Query Match 70.3%; Score 251; DB 12; Length 234;  
 Best Local Similarity 67.1%; Pred. No. 1.9e-20;  
 Matches 47; Conservative 13; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 TLWGEFQVDRSARGGCGTSDNAAAEALNASKEPALQIIREKIPEKYLFAFAALNSNL 60  
 Db 110 TIWGVFQIDGRSARGGCGTSDNAAAEALNSGTKEAAMRIVKEKLPKELFQYHNLSSDL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 DRIFMKAPNP 179  
 RESULT 12  
 Q92089 PRELIMINARY; PRT: 190 AA.  
 ID Q92089;  
 AC Q92089;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL AND COMPLETE CDS (FRAGMENT).  
 GN C1.  
 OS tobacco leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=67762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-YOKOHAMA3;  
 RA Ooi K., Ohshita S., Ishii I., Yahara T.;  
 RT "Molecular phylogeny of geminivirus infecting wild plants in Japan.";  
 RL J. Plant Res. 110:247-257(1997).  
 DR EMBL: AB001315; BAA34033.1; -  
 DR InterPro: IPR001191; Gemini\_AL1.  
 DR Pfam: PF00799; Gemini\_AL1; 1.

Search completed: January 3, 2002, 15:56:25

Job time: 1121 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:38:50 ; Search time 65.28 Seconds  
(without alignments)  
24.130 Million cell updates/sec

Title: US-09-289-346A-6  
Perfect score: 357  
Sequence: 1 TLWGEFQVDRSGRGCGT.....FAFAALNSNLDRIPTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	210	58.8	359	4	US-08-809-103B-2
2	210	58.8	359	4	US-08-809-103B-4
3	210	58.8	359	4	US-08-809-103B-6
4	210	58.8	359	4	US-08-809-103B-8
5	206	57.7	353	4	US-08-838-151A-44
6	206	57.7	353	4	US-08-838-151A-46
7	206	57.7	353	4	US-08-838-151A-49
8	206	57.7	353	4	US-08-838-151A-52
9	206	57.7	353	4	US-08-838-151A-55
10	204	57.1	361	4	US-08-838-151A-2
11	204	57.1	361	4	US-08-838-151A-4
12	204	57.1	361	4	US-08-838-151A-6
13	204	57.1	361	4	US-08-838-151A-8
14	191	53.5	357	4	US-08-838-151A-24
15	191	53.5	357	4	US-08-838-151A-27
16	191	53.5	357	4	US-08-838-151A-30
17	183	51.3	357	4	US-08-838-151A-20
18	64.5	18.1	1713	3	US-08-600-982-24
19	64.5	18.1	1713	5	PCT-US94-10261A-24
20	63.5	17.8	2161	1	US-07-745-206A-2
21	63.5	17.8	2161	1	US-08-455-543A-49
22	63.5	17.8	2161	1	US-08-455-543A-51
23	63.5	17.8	2161	2	US-08-223-305C-49
24	63.5	17.8	2161	2	US-08-223-305C-51
25	63.5	17.8	2161	2	US-08-311-363-2
26	60.5	16.9	131	4	US-08-838-151A-14
27	58	16.2	446	2	US-08-672-814D-11

28	58	16.2	446	4	US-09-333-696-11	Sequence 11, Appl
29	55.5	15.5	593	4	US-09-234-393-54	Sequence 54, Appl
30	55	15.4	374	4	US-09-091-405-2	Sequence 2, Appl
31	55	15.4	498	2	US-08-511-485-13	Sequence 13, Appl
32	54.5	15.3	335	2	US-09-014-969-15	Sequence 15, Appl
33	54	15.1	633	1	US-08-458-477A-5	Sequence 5, Appl
34	54	15.1	633	2	US-09-033-153-5	Sequence 5, Appl
35	54	15.1	633	4	US-09-325-430B-5	Sequence 5, Appl
36	54	15.1	765	2	US-08-663-112-2	Sequence 2, Appl
37	53.5	15.0	217	1	US-08-621-081A-23	Sequence 23, Appl
38	53.5	15.0	593	4	US-09-234-393-24	Sequence 24, Appl
39	53.5	15.0	593	4	US-09-234-393-50	Sequence 50, Appl
40	53.5	15.0	593	4	US-09-234-393-52	Sequence 52, Appl
41	53	14.8	266	2	US-07-857-224B-40	Sequence 40, Appl
42	53	14.8	374	2	US-08-928-692-51	Sequence 51, Appl
43	53	14.8	454	3	US-08-446-100-30	Sequence 30, Appl
44	52.5	14.7	217	1	US-08-621-081A-22	Sequence 22, Appl
45	52.5	14.7	382	2	US-08-477-451-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1  
US-08-809-103B-2  
; Sequence 2, Application US/08809103B  
; Patent No. 6133505  
; GENERAL INFORMATION:  
; APPLICANT: GRONENBORN, Bruno  
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT  
; TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: YOUNG & THOMPSON  
; STREET: 745 South 23rd Street  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/809,103B  
; FILING DATE: 17-MAR-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 94.11040  
; FILING DATE: 15-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/FR95/01192  
; FILING DATE: 15-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PATCH, Andrew J.  
; REGISTRATION NUMBER: 32,925  
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 521-2297  
; TELEFAX: (703) 685-0573  
; TELEX: 248425 EMBRON  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-809-103B-2

Query Match 58.8%; Score 210; DB 4; Length 359;  
Best Local Similarity 56.5%; Pred. No. 2.8e-20;





```
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94.11040
; FILING DATE: 15-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-809-103B-8

Query Match 58.8%; Score 210; DB 4; Length 359;
Best Local Similarity 56.5%; Pred. No. 2.8e-20;
Matches 39; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

Oy 2 LVWGEFQVDCRSARGGQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSLD 61
Db 111 LEWETFTQIDGHSARGGQQTANDAYAKAINAGSKQALDVIKELAPRDYVLHFHINSNLD 170

Oy 62 RIFDKTPEP 70
Db 171 KVFQVPAP 179

RESULT 5
US-08-838-151A-44
; Sequence 44, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,103B
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: SVS3801P0260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: SVS3801P0260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-838-151A-44

Query Match 57.7%; Score 206; DB 4; Length 353;
Best Local Similarity 57.1%; Pred. No. 9.5e-20;
Matches 40; Conservative 13; Mismatches 17; Indels 0; Gaps 0;

Oy 1 TLVWGEFQVDCRSARGGQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
Db 110 TIEWGQFQVDCRSARGGQGSANDSYAKALNADSIESTALTILKEQPKDYVLQHHNRSNL 169

Oy 61 DRIFDKTPEP 70
Db 170 ERIFVKVPEP 179

RESULT 6
US-08-838-151A-46
; Sequence 46, Application US/08838151A
; Patent No. 6291743
; GENERAL INFORMATION:
; APPLICANT: Stout, John T
; APPLICANT: Luu, Hang T
; APPLICANT: Maxwell, Douglas
; APPLICANT: Ahlquist, Paul
; APPLICANT: Hanson, Steve
; TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,151A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: SVS3801P0260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
```





US-08-838-151A-4

Query Match 57.1%; Score 204; DB 4; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 1.8e-19;  
 Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSARGCOTSDNDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
 Db 110 TIWGGDFQIDGRSARGGQSANDSYAKALNASSVQSALAVLREEQPKDFVLQNHNRSL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 ERIFAKAPEP 179

RESULT 12  
 US-08-838-151A-6  
 : Sequence 6, Application US/08838151A  
 : Patent No. 6291743  
 : GENERAL INFORMATION:  
 : APPLICANT: Stout, John T  
 : APPLICANT: Luu, Hang T  
 : APPLICANT: Maxwell, Douglas  
 : APPLICANT: Ahlquist, Paul  
 : APPLICANT: Hanson, Steve  
 : TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
 : NUMBER OF SEQUENCES: 63  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
 : STREET: Two Prudential Plaza, Suite 4700  
 : CITY: Chicago  
 : STATE: Illinois  
 : COUNTRY: U.S.A.  
 : ZIP: 60601  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/838,151A  
 : FILING DATE:  
 : CLASSIFICATION: 800  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Mueller, Lisa V  
 : REGISTRATION NUMBER: 38,978  
 : REFERENCE/DOCKET NUMBER: SVS3801P0260  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 312-616-5400  
 : TELEFAX: 312-616-5460  
 : INFORMATION FOR SEQ ID NO: 6:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 361 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-838-151A-6

Query Match 57.1%; Score 204; DB 4; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 1.8e-19;  
 Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSARGCOTSDNDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
 Db 110 TIWGGDFQIDGRSARGGQSANDSYAKALNASSVQSALAVLREEQPKDFVLQNHNRSL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 ERIFAKAPEP 179

RESULT 13  
 US-08-838-151A-8  
 : Sequence 8, Application US/08838151A  
 : Patent No. 6291743  
 : GENERAL INFORMATION:  
 : APPLICANT: Stout, John T  
 : APPLICANT: Luu, Hang T  
 : APPLICANT: Maxwell, Douglas  
 : APPLICANT: Ahlquist, Paul  
 : APPLICANT: Hanson, Steve  
 : TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
 : NUMBER OF SEQUENCES: 63  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
 : STREET: Two Prudential Plaza, Suite 4700  
 : CITY: Chicago  
 : STATE: Illinois  
 : COUNTRY: U.S.A.  
 : ZIP: 60601  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/838,151A  
 : FILING DATE:  
 : CLASSIFICATION: 800  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Mueller, Lisa V  
 : REGISTRATION NUMBER: 38,978  
 : REFERENCE/DOCKET NUMBER: SVS3801P0260  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 312-616-5400  
 : TELEFAX: 312-616-5460  
 : INFORMATION FOR SEQ ID NO: 8:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 361 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-838-151A-8

Query Match 57.1%; Score 204; DB 4; Length 361;  
 Best Local Similarity 54.3%; Pred. No. 1.8e-19;  
 Matches 38; Conservative 16; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVDRSARGCOTSDNDAAAEALNASSKEEALQIIREKIPKYLFAFAALNSNL 60  
 Db 110 TIWGGDFQIDGRSARGGQSANDSYAKALNASSVQSALAVLREEQPKDFVLQNHNRSL 169  
 QY 61 DRIFDKTPEP 70  
 Db 170 ERIFAKAPEP 179  
 RESULT 14  
 US-08-838-151A-24  
 : Sequence 24, Application US/08838151A  
 : Patent No. 6291743  
 : GENERAL INFORMATION:  
 : APPLICANT: Stout, John T  
 : APPLICANT: Luu, Hang T  
 : APPLICANT: Maxwell, Douglas  
 : APPLICANT: Ahlquist, Paul  
 : APPLICANT: Hanson, Steve  
 : TITLE OF INVENTION: Transgenic Plants Expressing Geminivirus  
 : NUMBER OF SEQUENCES: 63  
 : CORRESPONDENCE ADDRESS:

```

Query Match      53.5%; Score 191; DB 4; Length 357;
Best Local Similarity 62.3%; Pred. No. le-17;
Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY   4  WGEFQVDGRSARGCQT'SNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSLDRI 63
     : | : ||||| : ||| ||||| : ||| : ||| : | : | |||||
Db   111 FGVSQLDGRSARGGOQSANDAYAEALNSGSKSEALNILKERAPKDYLQLOFNLSNLDRI 170

QY   64 F 64
     . |
Db   171 F 171

```

Search completed: January 3, 2002, 15:38:50  
Job time: 227 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:40:10 ; Search time 72.79 seconds  
(without alignments)  
73.255 Million cell updates/sec

Title: US-09-289-346A-6  
Perfect score: 357  
Sequence: 1 TLVWGEFQVDRSARGCOT.....FAFAALNSLDRIFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	340	95.2	352	1 QOCVLI	ALI protein - toma
2	240	67.2	361	1 QOCVPT	ALI protein - toma
3	227	63.6	358	2 S07594	hypothetical prote
4	220	61.6	362	1 JQ1887	ALI protein - toma
5	217	60.8	359	2 S39211	gene C1 protein -
6	212	59.4	351	2 JQ2327	ALI protein - Indi
7	210	58.8	359	2 S22593	hypothetical prote
8	209	58.5	349	2 JQ2300	replicase - pepper
9	209	58.5	349	2 S31875	ALI protein - pepp
10	204	57.1	359	2 S39235	gene C1 protein -
11	203	56.9	355	1 QOCVW1	AVI protein - abut
12	200	56.0	358	1 JQ1870	ALI protein - toma
13	200	56.0	385	2 S28360	ALI protein - beet
14	193	54.1	360	2 S28360	replication-associ
15	191	53.5	357	1 QOCVCI	ALI protein - toma
16	118	33.1	131	2 S45059	AC1 protein (clone
17	104	29.1	347	1 QOCVSI	ALI protein - squa
18	67	18.8	447	2 T12544	hypothetical prote
19	64.5	18.1	587	2 JQ1419	Fc gamma (IgC) rec
20	64.5	18.1	1713	2 A53347	adhesive ligand ep
21	63.5	17.8	180	2 D84082	hypothetical prote
22	63.5	17.8	1610	2 A46227	voltage-dependent
23	63.5	17.8	1646	2 JH0422	voltage-dependent
24	63.5	17.8	2161	2 JH0564	calcium channel al
25	63.5	17.8	2181	2 A38198	calcium channel al
26	63.5	17.8	2203	2 T42742	voltage-dependent
27	63	17.6	316	2 C82085	conserved hypothe
28	61.5	17.2	392	2 T45290	acetyl-CoA C-acety
29	60.5	16.9	840	2 T36175	probable large ATP

30	60	16.8	269	2 S61633	YNT20 protein - ye
31	60	16.8	295	2 D42452	C1 protein - tobac
32	60	16.8	355	2 S49621	phytoene synthase
33	60	16.8	355	2 T50746	phytoene synthase
34	60	16.8	1888	2 T14273	zinc finger protei
35	59.5	16.7	299	2 B71967	probable peptidyl-
36	59.5	16.7	470	2 T46814	gamma-aminobutyrat
37	59.5	16.7	2137	2 T05244	hypothetical prote
38	59.5	16.7	4385	2 T29042	hypothetical prote
39	58.5	16.4	127	2 E86883	30S ribosomal prot
40	58.5	16.4	354	2 A75087	acetyl ornithine d
41	58.5	16.4	647	2 T26240	hypothetical prote
42	58.5	16.4	706	2 H71707	hypothetical prote
43	58.5	16.4	1203	2 T17415	mycelial surface a
44	58.5	16.4	1232	2 S40766	hypothetical prote
45	58.5	16.4	1501	2 C84512	probable retroelem

ALIGNMENTS

RESULT 1  
QOCVLI  
ALI protein - tomato golden mosaic virus  
C:Species: tomato golden mosaic virus  
A:Note: host Nicotiana sp. (tobacco)  
C:Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 08-Apr-1994  
C:Accession: A04170  
R:Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.  
EMBO J. 3, 2197-2205, 1984  
A:Title: Complete nucleotide sequence of the infectious cloned DNA components of toma  
A:Reference number: A04163  
A:Accession: A04170  
A:Molecule type: DNA  
A:Residues: 1-352 <HAM>  
C:Comment: The genome consists of two circular, single-stranded DNA components, DNA A  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 95.2%; Score 340; DB 1; Length 352;  
Best Local Similarity 95.7%; Pred. No. 4e-31;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGCOTSDAAAEALNASKEPALQIIREKIPEKYLFAFAALNSNL 60  
|||||  
Db 111 TLVWGEFQVDRSARGCOTSDAAAEALNASKEPALQIIREKIPEKYLFAFAALNSNL 170

QY 61 DRIFDKTPEP 70  
|||||  
Db 171 DRIFDKTPEP 180

RESULT 2  
QOCVPT  
ALI protein - potato yellow mosaic virus (isolate Venezuela)  
C:Species: potato yellow mosaic virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: J00364  
R:Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.  
J. Gen. Virol. 72, 1515-1520, 1991  
A:Title: The nucleotide sequence of the infectious cloned DNA components of potato ye  
A:Reference number: J00362; MUID:91311403  
A:Accession: J00364  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-361 <COU>  
A:Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus ALI protein

Query Match 67.2%; Score 240; DB 1; Length 361;  
Best Local Similarity 66.7%; Pred. No. 1e-19; Mismatches 11; Indels 0; Gaps 0;  
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 TLWGEFQIDGRSARGCGQTSNDAAAEALNASSKEALQIIREKIPEKYLFAFAALNSNL 60  
DB 110 TLWGLFQIDGRSARGCGQTVNDAAAEALNASSKEALQIIREKIPEKFLFOYHNLNSNL 169  
QY 61 DRIFDKTPE 69  
DB 170 DRIFWKAPE 178

RESULT 3  
S07594  
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)  
C:Species: cassava latent virus  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
C:Accession: S07594  
R:Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.  
Nucleic Acids Res. 18, 197-198, 1990  
A:Title: Nucleotide sequence of the infectious cloned DNA components of African cassava  
A:Reference number: S07594; MUID:90174930  
A:Accession: S07594  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-358 <MOR>  
A:Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376  
C:Genetics:  
A:Map position: segment DNAL  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 63.6%; Score 227; DB 2; Length 358;  
Best Local Similarity 60.0%; Pred. No. 3.1e-18; Mismatches 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 TLWGEFQIDGRSARGCGQTSNDAAAEALNASSKEALQIIREKIPEKYLFAFAALNSNL 60  
DB 109 TVWEGFQIDGRSARGCGQSDAYAKALNSGSKSEALNVIKELPKDVLQFVHNLNSNL 168  
QY 61 DRIFDKTPE 70  
DB 169 DRIFQEPAP 178

RESULT 4  
JQ1887  
ALL protein - tomato yellow leaf curl virus (strain Australia)  
N:Alternate names: Cl protein  
C:Species: tomato yellow leaf curl virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
C:Accession: JQ1887  
R:Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.  
J. Gen. Virol. 74, 147-151, 1993  
A:Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.  
A:Reference number: JQ1885; MUID:93139778  
A:Accession: JQ1887  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-362 <DR>  
A:Cross-references: GB:S33251  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 61.6%; Score 220; DB 1; Length 362;  
Best Local Similarity 52.9%; Pred. No. 2e-17; Mismatches 15; Indels 16; Gaps 1;  
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

QY 1 TLWGEFQIDGRSARGCGQTSNDAAAEALNASSKEALQIIREKIPEKYLFAFAALNSNL 60  
DB 111 TLWGLFQIDGRSARGCGQTVNDAAAEALNASSKEALQIIREKIPEKFLFOYHNLNSNL 169

DB 110 TLWGEFQIDGRSARGCGQSDAYAKALNSGSKSEALNVIKELPKDVLQFVHNLNSNL 169  
QY 61 DRI-----FDKTPPE 69  
DB 170 DRIFTPPLEVYVSPFLSSSFDVRPE 194

RESULT 5  
S39211  
gene Cl protein - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
C:Accession: S39211  
R:Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.  
submitted to the EMBL Data Library, August 1993  
A:Description: High similarity among the tomato yellow leaf curl virus isolates from  
A:Reference number: S39209  
A:Accession: S39211  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-359 <NOR>  
A:Cross-references: EMBL:Z23751; NID:g433655; PIDN:CAA81026.1; PID:g433658  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 60.8%; Score 217; DB 2; Length 359;  
Best Local Similarity 59.4%; Pred. No. 4.3e-17; Mismatches 41; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 2 LVWGEFQIDGRSARGCGQTSNDAAAEALNASSKEALQIIREKIPEKYLFAFAALNSNL 61  
DB 111 LEWGTQIDGRSARGCGQTANDAYAKAINAGSKSEALDKELAPRDYILHFNHNSNL 170  
QY 62 RIFDKTPEP 70  
DB 171 RVFQVPPAP 179

RESULT 6  
JQ2327  
ALL protein - Indian cassava mosaic virus  
N:Alternate names: replication-associated protein  
C:Species: Indian cassava mosaic virus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2327; S35883  
R:Hong, Y.G.; Robinson, D.J.; Harrison, B.D.  
J. Gen. Virol. 74, 2437-2443, 1993  
A:Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t  
A:Reference number: JQ2326; MUID:94065670  
A:Accession: JQ2327  
A:Molecule type: DNA  
A:Residues: 1-351 <HON>  
A:Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046  
C:Superfamily: tomato golden mosaic virus ALL protein

Query Match 59.4%; Score 212; DB 2; Length 351;  
Best Local Similarity 62.7%; Pred. No. 1.5e-16; Mismatches 8; Mismatches 17; Indels 0; Gaps 0;

QY 4 WGEFQIDGRSARGCGQTSNDAAAEALNASSKEALQIIREKIPEKYLFAFAALNSNL 63  
DB 113 WGTQIDGRSARGCGQSDAYAKALNSGSKSEALNVIKELPKDVLQFVHNLNSNL 172  
QY 64 FDKTPEP 70  
DB 173 FTKPPPP 179

RESULT 7  
S22593  
hypothetical protein C4 - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus



```

A:Reference number: A36214; MULD:1LU020984
A:Accession: A36214
A:Molecule type: DNA
A:Residues: 1-355 <FRI>
A:Cross-references: EMBL:X15983
C:Genetics:
A:Map position: segment A
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match          56.9%   Score 203;   DB 1;   Length 355;
Best Local Similarity 55.7%   Pred. No. 1.7e-15;
Matches 39; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

Ov 1 TLWGEFOVDCRSARGGCOTSNDAAEALNASSKEEALQITREKTPERYLFAFAALNSNL 60

```

Db 110 TAEMGEFQIDGRSARGGQQTANDSVAKALNAGDVQSAALNKLKEQPKDYVLQNHNRSL 169  
QY 61 DRIFDKTPEP 70  
Db 170 ERIFAKAPEP 179

## RESULT 12

AL1 protein - tomato mottle virus (Isolate Florida)  
A:Species: tomato mottle virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 07-May-1999  
C:Accession: J01870  
R:Abouzid, A.M.; Polston, J.E.; Hiebert, E.  
J. Gen. Virol. 73, 3225-3229, 1992  
A:Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from  
A:Reference number: J01870  
A:Accession: J01870  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-358 <ABO>  
A:Cross-References: GB:L14460  
C:Genetics:  
A:Map position: segment A  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 56.0%; Score 200; DB 1; Length 358;  
Best Local Similarity 52.9%; Pred. No. 3.7e-15;  
Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQIDGRSARGGQQTANDSVAKALNAGDVQSAALNKLKEQPKDYVLQNHNRSL 166  
Db 107 TIWGEFQIDGRSARGGQQTANDSVAKALNAGDVQSAALNKLKEQPKDYVLQNHNRSL 166  
QY 61 DRIFDKTPEP 70  
Db 167 ERIFAKAPEP 176

## RESULT 13

AL1 protein - beet curly top virus  
A:Species: beet curly top virus  
C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 20-Sep-1999  
C:Accession: S28360  
R:Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.  
EMBO J. 5, 1761-1767, 1986  
A:Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to  
A:Reference number: S28360  
A:Accession: S28360  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-385 <STA>  
A:Cross-References: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 56.0%; Score 200; DB 2; Length 385;  
Best Local Similarity 52.9%; Pred. No. 4e-15;  
Matches 37; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

QY 1 TLVWGEFQIDGRSARGGQQTANDSVAKALNAGDVQSAALNKLKEQPKDYVLQNHNRSL 160  
Db 137 TIWGEFQIDGRSARGGQQTANDSVAKALNAGDVQSAALNKLKEQPKDYVLQNHNRSL 196  
QY 61 DRIFDKTPEP 70  
Db 197 QKIFQRPDP 206

## RESULT 14

S59885  
replication-associated protein C1 - tomato yellow leaf curl virus  
C:Species: tomato yellow leaf curl virus  
C:Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Sep-1999  
C:Accession: S59885  
R:Hong, Y.; Harrison, B.D.  
submitted to the EMBL Data Library, February 1995  
A:Description: Nucleotide sequences from tomato leaf curl viruses from different coun  
d geminiviruses.  
A:Reference number: S58346  
A:Accession: S59885  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <HON>  
A:Cross-References: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 54.1%; Score 193; DB 2; Length 360;  
Best Local Similarity 57.6%; Pred. No. 2.3e-14;  
Matches 38; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 4 WGEFQVDCRSARGGQQTNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63  
Db 113 FGVFQIDGRSARGGQQTNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 172  
QY 64 FDKTPE 69  
Db 173 FTFPSAE 178

## RESULT 15

QOCVC1  
AL1 protein - tomato yellow leaf curl virus  
A:Alternate names: C1 protein  
C:Species: tomato yellow leaf curl virus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
C:Accession: D40779  
R:Navot, N.; Pichersky, E.; Zeldan, M.; Zamir, D.; Czosnek, H.  
Virology 185, 151-161, 1991  
A:Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin  
A:Reference number: A40779; MUID:92024070  
A:Accession: D40779  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-357 <NAV>  
A:Cross-References: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207  
C:Superfamily: tomato golden mosaic virus AL1 protein

Query Match 53.5%; Score 191; DB 1; Length 357;  
Best Local Similarity 62.3%; Pred. No. 3.9e-14;  
Matches 38; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 4 WGEFQVDCRSARGGQQTNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63  
Db 111 FGVFQIDGRSARGGQQTNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 170  
QY 64 F 64  
Db 171 F 171

Search completed: January 3, 2002, 15:40:10  
Job time: 306 sec



---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2002, 15:37:37 ; Search time 144.17 Seconds  
(without alignments)  
35.965 Million cell updates/sec

Title: US-09-289-346A-7  
Perfect score: 359  
Sequence: 1 TLVWGEFQVDRSARGCQT.....FOFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_1101.\*
- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
  - 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
  - 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
  - 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
  - 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
  - 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
  - 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
  - 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
  - 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
  - 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
  - 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
  - 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
  - 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
  - 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
  - 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
  - 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
  - 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
  - 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
  - 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
  - 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
  - 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
  - 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	359	100.0	70	AA18683	Mutant peptide der
2	342	95.3	70	AA18677	Peptide fragment f
3	342	95.3	356	AA18687	Amino acid sequenc
4	334	93.0	70	AA18685	Mutant peptide der
5	331	92.2	70	AA18688	Mutant peptide der
6	330	91.9	70	AA18692	Mutant peptide der
7	328	91.4	70	AA18684	Mutant peptide der
8	328	91.4	70	AA18690	Mutant peptide der
9	327	91.1	70	AA18678	Mutant peptide der
10	327	91.1	70	AA18686	Mutant peptide der
11	326	90.8	70	AA18689	Mutant peptide der

12	324	90.3	70	21	AA18680	Mutant peptide der
13	324	90.3	70	21	AA18691	Mutant peptide der
14	321	89.4	70	21	AA18681	Mutant peptide der
15	318	88.6	70	21	AA18682	Mutant peptide der
16	310	86.4	70	21	AA18679	Mutant peptide der
17	213	59.3	353	18	AAW34338	Bean golden mosaic
18	213	59.3	353	18	AAW34332	Bean golden mosaic
19	213	59.3	353	18	AAW34333	Bean golden mosaic
20	213	59.3	353	18	AAW34334	Bean golden mosaic
21	213	59.3	353	18	AAW34335	Bean golden mosaic
22	211	58.8	353	8	AA170407	ORF 4 gene product
23	211	58.8	359	17	AA186870	Sardinian tomato y
24	211	58.8	359	17	AA186871	Sardinian tomato y
25	211	58.8	359	17	AA186872	Sardinian tomato y
26	211	58.8	361	18	AAW34336	Tomato mottle viru
27	211	58.8	361	18	AAW34324	Tomato mottle viru
28	211	58.8	361	18	AAW34325	Tomato mottle viru
29	211	58.8	361	18	AAW34326	Tomato mottle viru
30	208.5	58.1	361	8	AA170562	Product of ORF 4 f
31	194	54.0	362	19	AAW56495	Tobacco leaf curl
32	193	53.8	357	18	AAW34329	Tomato yellow leaf
33	193	53.8	357	18	AAW34330	Tomato yellow leaf
34	193	53.8	357	18	AAW34331	Tomato yellow leaf
35	185	51.5	357	18	AAW34337	Tomato yellow leaf
36	68	18.9	447	21	AA199659	Human GTPase Assoc
37	68	18.9	447	21	AA197089	Human secreted pro
38	68	18.9	447	21	AA166744	Membrane-bound pro
39	68	18.9	447	21	AA150947	Human adult aorta
40	68	18.9	447	22	AAE06066	Human gene 26 enco
41	68	18.9	447	22	AAE65267	Human PRO1125 (UNQ
42	68	18.9	456	21	AA187190	Human secreted pro
43	68	18.9	456	22	AAE06167	Human gene 26 enco
44	63.5	17.7	619	13	AA127651	Human calcium chan
45	63.5	17.7	2161	14	AA133545	Sequence of the al

## ALIGNMENTS

RESULT	1	
AA18683	ID	AA18683 standard; peptide; 70 AA.
XX	AC	AA18683;
XX	AC	AA18683;
XX	AC	AA18683;
DT	22-JAN-2001	(first entry)
XX	XX	Mutant peptide derived from amino acids 110-179 of Rep (A11) protein.
DE	XX	Geminivirus; replication protein; Rep protein; A11; transgenic plant;
XX	XX	ribosome binding region; resistance; geminivirus infection.
KW	XX	Synthetic.
XX	OS	Tomato golden mosaic virus.
OS	XX	Tomato golden mosaic virus.
XX	XX	Key Location/Qualifiers
FT	FT	Misc-difference 59 /note= "wild type residue replaced with Ala"
FT	FT	Misc-difference 61 /note= "wild type residue replaced with Ala"
FT	FT	Misc-difference 62 /note= "wild type residue replaced with Ala"
FT	FT	Misc-difference 62 /note= "wild type residue replaced with Ala"
XX	XX	WO200054573-A1.
FN	XX	21-SEP-2000.
XX	XX	15-MAR-2000; 2000WO-US06759.
XX	XX	18-MAR-1999; 99US-0125004.
PR	XX	09-APR-1999; 99US-0289346.
XX	XX	(UYN:-) UNIV NORTH CAROLINA STATE.
XX	XX	

XX  
PI Hanley-Bowdoin L, Orozco BM, Kong L;  
XX WPI; 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
PS Claim 53; Page 45; 73pp; English.  
XX  
XX The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyzes cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA;  
  
Query Match 100.0%; Score 359; DB 21; Length 70;  
Best Local Similarity 100.0%; Pred. No. 4.9e-39;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TLVWGEFQVDRSGCGCOTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLSAL 60  
Db 1 TLVWGEFQVDRSGCGCOTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLSAL 60  
QY 61 AAIFDKTPEP 70  
Db 61 aaifdktp 70  
  
RESULT 2  
AAB18677  
ID AAB18677 standard; peptide: 70 AA.  
XX  
XX AAB18677;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Peptide fragment from Rep protein of TGMV (amino acids 110-179).  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Tomato golden mosaic virus.  
XX  
XX WO200054573-A1.  
PN  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
PI WPI; 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection

PT Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
PS Disclosure; Page 18; 73pp; English.  
XX  
XX The present sequence is derived from a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyzes cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
XX  
SQ Sequence 70 AA;  
  
Query Match 95.3%; Score 342; DB 21; Length 70;  
Best Local Similarity 95.7%; Pred. No. 7.7e-37;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 TLVWGEFQVDRSGCGCOTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLSAL 60  
Db 1 TLVWGEFQVDRSGCGCOTSDNDAAEALNASSKEEALQIIREKIPEKYLFOFHNLSAL 60  
QY 61 AAIFDKTPEP 70  
Db 61 drifdktp 70  
  
RESULT 3  
AAB18687  
ID AAB18687 standard; peptide: 356 AA.  
XX  
XX AAB18687;  
XX  
XX 22-JAN-2001 (first entry)  
XX  
XX Amino acid sequence of a geminivirus replication protein of TGMV.  
XX  
XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
XX Tomato golden mosaic virus.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 354  
FT /note= "unspecified amino acid"  
XX  
XX WO200054573-A1.  
PN  
XX  
XX 21-SEP-2000.  
XX  
XX 15-MAR-2000; 2000WO-US06759.  
XX  
XX 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Hanley-Bowdoin L, Orozco BM, Kong L;  
PI WPI; 2000-618851/59.  
XX  
XX Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
PS Disclosure; Page 47-48; 73pp; English.  
XX

CC The present sequence represents a geminivirus replication (Rep)  
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,  
CC catalyses cleavage and ligation of single-stranded DNA, and interacts  
CC with other viral and host proteins. Mutants of the AL1 protein are used  
CC to produce transgenic plants. The mutation in AL1 is present in a  
CC ribosome binding region, and expression of mutant AL1 protein imparts  
CC increased resistance to geminivirus infection in the plant. Mutant AL1  
CC proteins are useful for producing plants having increased resistance or  
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,  
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl  
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato  
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,  
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or  
CC beet curly top virus.  
XX

SQ Sequence 356 AA;

Query Match 95.3%; Score 342; DB 21; Length 356;  
Best Local Similarity 95.7%; Pred. No. 6.3e-36;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGGQTSNDAAAEALNASSKEALQIIREKIPEKYLFGFHNLSAL 60  
|||||  
Db 110 tlwgefvgdrgsarggcqtsndaaaealnasskealqirekipekylfgfhnlnsl 169

QY 61 AAIFDKTPEP 70  
|||||  
Db 170 drifdktp 179

#### RESULT 4

AAB18685  
ID AAB18685 standard; peptide; 70 AA.

XX  
AC AAB18685;

XX 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

OS Tomato golden mosaic virus.

XX Key Location/Qualifiers

FH Misc-difference 10 /note= "wild type residue replaced with Ala"  
FT

XX WO200054573-A1.

PN 21-SEP-2000.

PD 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.

PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
PT  
XX  
PS Claim 53; Page 46; 73pp; English.  
XX

CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX

SQ Sequence 70 AA;

Query Match 93.0%; Score 334; DB 21; Length 70;  
Best Local Similarity 94.3%; Pred. No. 8.4e-36;  
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TLVWGEFQVDRSARGGQTSNDAAAEALNASSKEALQIIREKIPEKYLFGFHNLSAL 60  
|||||  
Db 1 tlwgefvgdrgsarggcqtsndaaaealnasskealqirekipekylfgfhnlnsl 60

QY 61 AAIFDKTPEP 70  
|||||  
Db 61 drifdktp 70

#### RESULT 5

AAB18688  
ID AAB18688 standard; peptide; 70 AA.

XX  
AC AAB18688;

XX 22-JAN-2001 (first entry)

DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.

XX Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.

XX Synthetic.

OS Tomato golden mosaic virus.

XX Key Location/Qualifiers

FH Misc-difference 19 /note= "wild type residue replaced with Ala"  
FT

FT Misc-difference 20 /note= "wild type residue replaced with Ala"  
FT

XX WO200054573-A1.

PN 21-SEP-2000.

XX 15-MAR-2000; 2000WO-US06759.

XX 18-MAR-1999; 99US-0125004.

PR 09-APR-1999; 99US-0289346.

XX (UYNC-) UNIV NORTH CAROLINA STATE.

XX Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.

XX Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
PS Disclosure; Page 48; 73pp; English.  
XX  
CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA;  
  
Query Match 92.2%; Score 331; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 2.1e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLWGEFQVDRSARGCGTSDNDAAEALNASSKEALQIIREKIPEKYLFGFHNLSAL 60  
|||||  
Db 1 tlwgefqvdrsgargcgtsdaasndaaaealnasskeaalqirekipekylfgfhnlnsl 60  
|||||  
QY 61 AAIFDKTPEP 70  
|||||  
Db 61 drifdktppep 70  
  
RESULT 6  
AAB18692  
ID AAB18692 standard; peptide; 70 AA.  
XX  
AC AAB18692;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 66 /note= "wild type residue replaced with Ala"  
FT Misc-difference 69 /note= "wild type residue replaced with Ala"  
FT  
XX  
XX  
PN WO200054573-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 15-MAR-2000; 2000WO-US06759.  
XX  
PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
PA (UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
PI Hanley-Bowdoin L, Orozco BM, Kong L;

XX WPI; 2000-618851/59.  
XX  
PT Transgenic plants with increased resistance to geminivirus infection  
PT comprise a nucleic acid construct containing a nucleic acid sequence  
PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
XX  
PS Disclosure; Page 50; 73pp; English.  
XX  
CC The present sequence represents a mutant peptide, derived from a  
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
CC protein are used to produce transgenic plants. The mutation in AL1 is  
CC present in a ribosome binding region, and expression of mutant AL1  
CC protein imparts increased resistance to geminivirus infection in the  
CC plant. Mutant AL1 proteins are useful for producing plants having  
CC increased resistance or reduced sensitivity to a geminivirus such as  
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
CC virus, cotton leaf curl virus or beet curly top virus.  
XX  
SQ Sequence 70 AA;  
  
Query Match 91.9%; Score 330; DB 21; Length 70;  
Best Local Similarity 92.9%; Pred. No. 2.8e-35;  
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TLWGEFQVDRSARGCGTSDNDAAEALNASSKEALQIIREKIPEKYLFGFHNLSAL 60  
|||||  
Db 1 tlwgefqvdrsgargcgtsdaaeealnasskealqirekipekylfgfhnlnsl 60  
|||||  
QY 61 AAIFDKTPEP 70  
|||||  
Db 61 drifdktppep 70  
  
RESULT 7  
AAB18684  
ID AAB18684 standard; peptide; 70 AA.  
XX  
AC AAB18684;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
XX  
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
KW ribosome binding region; resistance; geminivirus infection.  
XX  
OS Synthetic.  
OS Tomato golden mosaic virus.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 7 /note= "wild type residue replaced with Ala"  
FT Misc-difference 8 /note= "wild type residue replaced with Ala"  
FT  
XX  
XX  
PN WO200054573-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 15-MAR-2000; 2000WO-US06759.  
XX  
PR 18-MAR-1999; 99US-0125004.  
PR 09-APR-1999; 99US-0289346.  
XX  
PA (UYNC-) UNIV NORTH CAROLINA STATE.  
XX



```
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX DR WPI; 2000-618851/59.
XX PT Transgenic plants with increased resistance to geminivirus infection
XX PT comprise a nucleic acid construct containing a nucleic acid sequence
XX PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX PS
XX PS Claim 52; Page 45; 73pp; English.
XX CC The present sequence represents a mutant peptide, derived from a
XX CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX CC protein are used to produce transgenic plants. The mutation in AL1 is
XX CC present in a ribosome binding region, and expression of mutant AL1
XX CC protein imparts increased resistance to geminivirus infection in the
XX CC plant. Mutant AL1 proteins are useful for producing plants having
XX CC increased resistance or reduced sensitivity to a geminivirus such as
XX CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX CC virus, cotton leaf curl virus or beet curly top virus.
XX SQ Sequence 70 AA;
Query Match 91.4%; Score 328; DR 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 5e-35;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 TLVWGEFQVDRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSAL 60
Db 1 TLVWGEAVDGRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPEKYLQFHNLSNL 60
Qy 61 AAIFDKTPPEP 70
Db 61 drifdktppep 70
RESULT 8
AAB18690
ID AAB18690 standard; peptide; 70 AA.
XX AC AAB18690;
XX DT 22-JAN-2001 (first entry)
XX DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX KW ribosome binding region; resistance; geminivirus infection.
XX OS Synthetic.
XX OS Tomato golden mosaic virus.
XX FH Key Location/Qualifiers
XX FT Misc-difference 27 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 30 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 30 /note= "wild type residue replaced with Ala"
XX FT
XX PN WO200054573-Al.
XX PD 21-SEP-2000.
XX PF 15-MAR-2000; 2000WO-US06759.
XX PR 18-MAR-1999; 99US-0125004.
XX PR 09-APR-1999; 99US-0289346.
```

```
XX PA (UYN-) UNIV NORTH CAROLINA STATE.
XX PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX DR WPI; 2000-618851/59.
XX PT Transgenic plants with increased resistance to geminivirus infection
XX PT comprise a nucleic acid construct containing a nucleic acid sequence
XX PT encoding a mutant AL1 protein with a mutation in the Rb binding region
XX PS
XX PS Disclosure; Page 49; 73pp; English.
XX CC The present sequence represents a mutant peptide, derived from a
XX CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
XX CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
XX CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
XX CC protein are used to produce transgenic plants. The mutation in AL1 is
XX CC present in a ribosome binding region, and expression of mutant AL1
XX CC protein imparts increased resistance to geminivirus infection in the
XX CC plant. Mutant AL1 proteins are useful for producing plants having
XX CC increased resistance or reduced sensitivity to a geminivirus such as
XX CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
XX CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
XX CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
XX CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
XX CC virus, cotton leaf curl virus or beet curly top virus.
XX SQ Sequence 70 AA;
Query Match 91.4%; Score 328; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 5e-35;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 TLVWGEFQVDRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPEKYLFOFHNLNSAL 60
Db 1 TLVWGEFQVDRSGRGCGQTSNDAAAEALNASSKEEALQIIREKIPEKYLQFHNLSNL 60
Qy 61 AAIFDKTPPEP 70
Db 61 drifdktppep 70
RESULT 9
AAB18678
ID AAB18678 standard; peptide; 70 AA.
XX AC AAB18678;
XX DT 22-JAN-2001 (first entry)
XX DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX DE Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
XX KW ribosome binding region; resistance; geminivirus infection.
XX OS Synthetic.
XX OS Tomato golden mosaic virus.
XX FH Key Location/Qualifiers
XX FT Misc-difference 12 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 13 /note= "wild type residue replaced with Ala"
XX FT Misc-difference 15 /note= "wild type residue replaced with Ala"
XX FT
XX PN WO200054573-Al.
XX PD 21-SEP-2000.
XX PR
```





OS Tomato golden mosaic virus.  
 XX Key Location/Qualifiers  
 FH Misc-difference 34 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 35 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 36 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 PN 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 XX  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI; 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Disclosure: Page 49; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 Query Match 90.3%; Score 324; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 1,7e-34;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVGRSARGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFGFHNLSAL 60  
 Db 1 tlvwgefvdgrsarggcqtsndaaaealnasskeaalqilirekipekylfgfhnlnsnl 60  
 QY 61 AAIFDKTPEP 70  
 Db 61 drifdktppep 70  
 RESULT 14  
 AAB18681  
 ID AAB18681 standard; peptide; 70 AA.  
 XX AAB18681;  
 XX  
 XX 22-JAN-2001 (first entry)  
 DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.  
 XX

KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;  
 KW ribosome binding region; resistance; geminivirus infection.  
 XX Synthetic.  
 OS Tomato golden mosaic virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 47 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 48 /note= "wild type residue replaced with Ala"  
 FT Misc-difference 49 /note= "wild type residue replaced with Ala"  
 FT  
 XX WO200054573-A1.  
 PN 21-SEP-2000.  
 XX  
 XX 15-MAR-2000; 2000WO-US06759.  
 XX  
 XX 18-MAR-1999; 99US-0125004.  
 PR 09-APR-1999; 99US-0289346.  
 XX  
 XX (UYNC-) UNIV NORTH CAROLINA STATE.  
 PA  
 XX Hanley-Bowdoin L, Orozco BM, Kong L;  
 PI WPI; 2000-618851/59.  
 DR  
 XX Transgenic plants with increased resistance to geminivirus infection  
 PT comprise a nucleic acid construct containing a nucleic acid sequence  
 PT encoding a mutant AL1 protein with a mutation in the Rb binding region  
 PT  
 XX  
 PS Claim 52; Page 44; 73pp; English.  
 XX  
 CC The present sequence represents a mutant peptide, derived from a  
 CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds  
 CC double-stranded DNA, catalyses cleavage and ligation of single-stranded  
 CC DNA, and interacts with other viral and host proteins. Mutants of the AL1  
 CC protein are used to produce transgenic plants. The mutation in AL1 is  
 CC present in a ribosome binding region, and expression of mutant AL1  
 CC protein imparts increased resistance to geminivirus infection in the  
 CC plant. Mutant AL1 proteins are useful for producing plants having  
 CC increased resistance or reduced sensitivity to a geminivirus such as  
 CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl  
 CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian  
 CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic  
 CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper  
 CC virus, cotton leaf curl virus or beet curly top virus.  
 XX  
 SQ Sequence 70 AA;  
 Query Match 89.4%; Score 321; DB 21; Length 70;  
 Best Local Similarity 91.4%; Pred. No. 4,1e-34;  
 Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 TLVWGEFQVGRSARGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFGFHNLSAL 60  
 Db 1 tlvwgefvdgrsarggcqtsndaaaealnasskeaalqilirekipekylfgfhnlnsnl 60  
 QY 61 AAIFDKTPEP 70  
 Db 61 drifdktppep 70  
 RESULT 15  
 AAB18682  
 ID AAB18682 standard; peptide; 70 AA.  
 XX AAB18682;  
 XX

Job time: 154 sec

DT	22-JAN-2001	(first entry)
XX		
DE	Mutant peptide derived from amino acids.110-179 of Rep (AL1) protein.	
XX		
KW	Geminivirus; replication protein; Rep protein; AL1; transgenic plant;	
KW	ribosome binding region; resistance; geminivirus infection.	
XX		
OS	Synthetic.	
OS	Tomato golden mosaic virus.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 52	/note= "wild type residue replaced with Ala"
FT	Misc-difference 54	/note= "wild type residue replaced with Ala"
FT	Misc-difference 55	/note= "wild type residue replaced with Ala"
FT		
XX	WO200054573-A1.	
PN		
XX	21-SEP-2000.	
PD		
XX	15-MAR-2000; 2000WO-US06759.	
XX		
XX	18-MAR-1999; 99US-0125004.	
PR	09-APR-1999; 99US-0289346.	
PR		
XX	(UYNC-) UNIV NORTH CAROLINA STATE.	
PA		
XX		
PI	Hanley-Bowdoin L, Orozco BM, Kong L;	
XX	WPI; 2000-618851/59.	
DR		
XX		
PT	Transgenic plants with increased resistance to geminivirus infection	
PT	comprise a nucleic acid construct containing a nucleic acid sequence	
PT	encoding a mutant AL1 protein with a mutation in the Rb binding region	
PT		
XX		
PS	Claim 53; Page 44-45; 73pp; English.	
XX		
CC	The present sequence represents a mutant peptide, derived from a	
CC	geminivirus replication (Rep) protein, also known as AL1. AL1 binds	
CC	double-stranded DNA, catalyses cleavage and ligation of single-stranded	
CC	DNA, and interacts with other viral and host proteins. Mutants of the AL1	
CC	protein are used to produce transgenic plants. The mutation in AL1 is	
CC	present in a ribosome binding region, and expression of mutant AL1	
CC	protein imparts increased resistance to geminivirus infection in the	
CC	plant. Mutant AL1 proteins are useful for producing plants having	
CC	increased resistance or reduced sensitivity to a geminivirus such as	
CC	tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl	
CC	virus, tomato leaf curl virus, African cassava mosaic virus, Indian	
CC	cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic	
CC	virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper	
CC	virus, cotton leaf curl virus or beet curly top virus.	
XX		
SQ	Sequence 70 AA:	
	Query Match 88.6%; Score 318; DB 21; Length 70;	
	Best Local Similarity 91.4%; Pred. No. 9.9e-34;	
	Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
Qy	1 TLVWGFQVDRGARGGQOTSNDAAAEALNASSKEEALQITREKIPEKYLFOHNLNSAL 60	
Db	1 tlvwgfqvdrgarggcqtsndaaaealnasskeeaqlirekipekylfafaalnsnl 60	
Qy	61 AAIFDKTPEP 70	
Db	61 drifdktppep 70	

